

THE INTEGRATION OF LINGUISTIC AND NON-LINGUISTIC INFORMATION
IN SECOND LANGUAGE SENTENCE PROCESSING

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF
THE UNIVERSITY OF HAWAII AT MĀNOA
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY
IN
SECOND LANGUAGE STUDIES

May 2017

By
Hyunah Ahn

Dissertation Committee:
William O’Grady, Chairperson
Victoria Anderson
Richard Day
Nicole Ziegler
Min Liu

Keywords: L2 psycholinguistics, information integration, English articles, real-world knowledge

To all who speak more than one language

ACKNOWLEDGEMENT

While working on my degree, I have owed so much to the members of my dissertation committee, other faculty and staff members in and outside the Department of Second Language Studies (SLS), and friends and family. I would never have come this far without everyone who was there for me to walk this path and finish this leg of journey in my academic career that is only beginning.

I would first like to express my deepest gratitude to my dissertation committee: Dr. William O'Grady, Dr. Victoria Anderson, Dr. Richard Day, Dr. Nicole Ziegler, and Dr. Min Liu. This dissertation would not have been possible without the insightful feedback and guidance I had privilege to receive from the members of the committee on every step of the way.

I would also like to thank professors who are not on my committee but have helped me tremendously. Coursework and guidance that was made available to me through Dr. James Dean Brown and Dr. Patricia Donegan extended the depth and breadth of my research agenda. I appreciate all the administrative and moral support throughout the entire process of my dissertation that I have gotten from Dr. Graham Crookes, the department chair of SLS and Dr. Gabriel Kasper, the graduate chair of SLS. I would also like to acknowledge Dr. Mijeong Song at Seoul National University (SNU) for her encouragement and support both when I was studying under her supervision at SNU and when I collected data for this dissertation in Seoul.

Conducting this research was possible thanks to nearly 500 participants from Honolulu and Seoul and institutions that provided financial support—a *Language Learning* dissertation grant and an East-West Center research grant. I thank Raina Heaton who worked with me to record all the auditory stimuli in the dissertation. Minsun Bong and Sarah Hyeyeon Lee helped me with data collection at SNU, which I feel greatly indebted to.

While working on my degree, I have served the SLS department as a graduate assistant mostly teaching undergraduate courses. Working with students from diverse backgrounds would have been far greater a challenge without the support and guidance from the undergraduate program coordinator of SLS, Mr. Kenton Harsch. As anyone who has worked with him would agree, I feel lucky that Kenny was my supervisor at work.

Some of my time as a graduate assistant was spent in the Language Analysis and Experimentation (LAE) Laboratories. The late Mr. Kurt Brunner, the lab technician of the LAE

labs deserves acknowledgement for his dedicated service and kindness. Although he was not with us anymore by the time I started this dissertation, I would not have learned about working in the labs and using various lab equipment and software programs as much as I did had it not been for his help and guidance.

Writing a dissertation and finishing a degree takes much more than merely seeking academic and professional training and guidance. Special thanks go to Dr. William O'Grady and Dr. Victoria Anderson for their patience, empathy, and unwavering trust they have shown me. When I was crestfallen and intimidated for academic or non-academic challenges, the two professors never stopped showing the deepest empathy or encouraging me to keep pushing forward. I will make my best efforts to emulate them for their excellence in both research and mentoring.

Around the end of my third year, I realized life exists outside the labs. I thank all my friends who were there for me sharing laughs and tears. My buddies in SLS: Angela Haeusler, Bonnie Sylwester, Hanbyul Jung, Kyaesung Park, Aya Takeda, and Mari Miyao; the Korean gang in linguistics: Jonny Kim, Onsoon Lee, ChaeEun Kim, Sunju Kim, and Jinsun Choe; and finally, the linguists in crime: James Grama, Raina Heaton, Joelle Kirtley, Melody Ross, and Claire Stabile.

Lastly, I would like to thank my family in Korea for their love and care and for all the wait and prayers: my parents Pyongkuk Ahn and Chungye Cho, and my extended family who filled my empty spot for my parents when they needed me.

ABSTRACT

This dissertation investigates the integration of linguistic and non-linguistic information in the course of second language (L2) sentence processing. Having to process multiple sources of information simultaneously has been claimed to pose the greatest challenges in L2 learning (Clahsen & Felser, 2006; Sorace, 2011; Sorace & Filiaci, 2006), and research on the issue has the potential to shed light on the nature of learning in different populations, which, in turn, could contribute to a better understanding of the role of processing in language acquisition (O’Grady, 2015; Phillips & Ehrenhofer, 2015).

The two sources of information, linguistic and non-linguistic knowledge, were operationalized by the definiteness distinction of English articles and real-world knowledge, respectively. To compare how different sources of information are integrated by L1 and L2 speakers, it was first necessary to determine whether both types of information were shared by native (L1) and second language (L2) speakers.

The first experiment, implemented via a self-paced reading task, examined whether L1 and L2 speakers are sensitive to the mapping between definite noun phrases (NPs) and unique referents. It was shown that both populations exhibit this pattern of mapping, but that L2 speakers’ sensitivity to the relationship appeared one region later compared to L1 speakers.

The second experiment, via a referent prediction task, shows that both L1 and advanced L2 speakers predict a unique referent at the cue of a definite article. What is noteworthy in this experiment is the behavior of intermediate L2 speakers, who predict a unique referent in response to an indefinite article numerically more often than to a definite article.

The third experiment was an online norming survey to confirm that both L1 and L2 speakers have the same real-world knowledge. The types of world knowledge studied consisted of associations between two referents—for example, a doctor and a stethoscope, or a basketball player and a basketball. L1 speakers and L2 speakers of all proficiency levels were found to share the same real-world knowledge.

The final experiment tested how the two sources of information were integrated incrementally online. A referent identification task measured the reaction time to stimuli when linguistic and non-linguistic information pointed to either the same referent or different referents. The results showed that L1 speakers integrate both linguistic and non-linguistic information incrementally and use both types of information to predict referents yet to be mentioned, but that L2 speakers did not use linguistic information in a native-like manner when non-linguistic information alone was sufficient to predict upcoming linguistic material.

The findings suggest that non-linguistic information, operationalized as real-world knowledge in the current research, could be the key to accounting for certain differences between L1 and L2 development. Such findings have important implications for issues in both psycholinguistics and language acquisition research. In particular, L2 speakers' reliance on real-world knowledge could be interpreted as an effort to minimize processing cost. By focusing their limited cognitive resources on an information source (world knowledge) that is acquired earlier than the relevant components of L2 grammar, is more familiar, and is thus easier-to-process, L2 speakers can maximize their processing efficiency. In contrast, it would not be as efficient to divide attentional resources over multiple sources of information. The dissertation concludes with suggestions for future research that compares adult L2 speakers and L1 children in terms of information integration to better understand what sets L2 acquisition apart from L1 acquisition.

TABLE OF CONTENTS

Acknowledgement	iv
Abstract	vi
Table of Contents	viii
List of Tables	xi
List of Figures	xiii
Chapter 1. Introduction	1
Chapter 2. Background	
2.1 Investigating processing to explain learning.....	5
2.2 Definiteness in L1 and L2 acquisition and processing research	7
2.3 The semantics of definiteness	11
2.4 The role of real-world knowledge in L1 and L2 sentence processing	
2.4.1 Studies on real-world knowledge in L1 sentence processing	14
2.4.2 Studies on real-world knowledge in L2 sentence processing	23
2.5 Research questions	26
Chapter 3. Experiment 1: Processing of definiteness as uniqueness	
3.1 Introduction	28
3.2 Background	28
3.2.1 L1 research on definiteness processing	29
3.2.2 Research on L2 English article use	35
3.3 Method	
3.3.1 Self-paced reading (SPR)	42
3.3.2 Proficiency	44
3.3.3 Participants	44
3.3.4 Stimuli	45
3.3.5 Procedure	46
3.4 Analysis	
3.4.1 Mixed effects modeling	48
3.4.2 Assumptions of linear regression	49
3.4.3 Residual Analysis	53

3.5	Results	55
3.5.1	L1 speakers	56
3.5.2	Advanced L2 speakers	61
3.6	Discussion	66
Chapter 4.	Experiment 2: Definiteness as a predictive cue for unique identifiability	
4.1	Introduction	71
4.2	Background	71
4.3	Method	
4.3.1	Referent Prediction Task	75
4.3.2	Proficiency and participants	78
4.3.3	Stimuli	79
4.3.4	Procedure	79
4.4	Analysis	81
4.5	Results	85
4.5.1	L1 speakers	86
4.5.2	Advanced L2 speakers	88
4.5.3	Intermediate L2 speakers	89
4.5.4	Interaction of article and speaker group in group comparisons	91
4.6	discussion	96
Chapter 5.	The integration of linguistic and non-linguistic information in L2 processing	
5.1	Introduction	99
5.2	Background	100
5.3	Experiment 3a: World knowledge norming	105
5.4	Experiment 3b: Integration	
5.4.1	Method	
5.4.1.1	Participants and proficiency	106
5.4.1.2	Stimuli	107
5.4.1.3	Procedure	111
5.4.2	Analysis.....	111
5.4.2.1	Log transformation	112
5.4.2.2	Residual reaction time and linear mixed effects regression	113
5.4.3	Results	116
5.4.3.1	L1 speakers	117
5.4.3.2	Advanced L2 speakers	124
5.4.3.3	Intermediate L2 speakers	129
5.5	Discussion	135

Chapter 6. Conclusion	
6.1 Summary	139
6.2 Theoretical implications	142
6.3 Pedagogical implications	143
6.4 Future research	146
Appendices	
Appendix 1a: C-test for proficiency	149
Appendix 1b: Experiment 1– Critical items	150
Appendix 1c: Experiment 1– Filler items	154
Appendix 2a: Experiment 2–Critical items	158
Appendix 2b: Experiment 2–Filler samples	162
Appendix 3a: Experiment 3–Critical items	163
Appendix 3b–Experiment 3: Filler samples	168
References	169

List of Tables

Table 3.1	Definiteness distinction in English
Table 3.2	Specificity distinction in Samoan
Table 3.3	Fluctuation between definiteness and specificity in Interlanguage grammar
Table 3.4	Revised <i>le/se</i> distinction in Samoan
Table 3.5	Participant demographics
Table 3.6	Descriptive statistics of reading time data before and after log transformation
Table 3.7	Linear mixed effects regression output for log RT residual analysis in L1 data
Table 3.8	Linear mixed effects regression output for log RT residual analysis in L2 data
Table 3.9	The lmer output for the interaction between article and context in L1
Table 3.10	The lmer output for the interaction between article and context in advanced L2
Table 4.1	Participant demographics
Table 4.2	The glmer output for the main effect of article in L1 speakers
Table 4.3	The glmer output for the main effect of article in advanced L2 speakers
Table 4.4	The glmer output for the main effect of article in intermediate L2 speakers
Table 4.5	The glmer output for the interaction of article and group in L1 and adv. L2
Table 4.6	The glmer output for the interaction of article and group in L1 and inter. L2
Table 4.7	The glmer output for the interaction of article and group between adv. and int. L2
Table 4.8	The glmer output for the interaction of article and group b/w inter. L2 and the rest
Table 5.1	Participant demographics
Table 5.2	Descriptive statistics for reaction time data before and after log transformation
Table 5.3	Linear mixed effects regression output for log RT residual analysis in L1

Table 5.4	Linear mixed effects regression output for log RT residual analysis in adv. L2
Table 5.5	Linear mixed effects regression output for log RT residual analysis in inter. L2
Table 5.6	Main and interaction effects of grammatical knowledge and world knowledge in L1
Table 5.7	The log RT residuals by condition in L1
Table 5.8	Main and interaction effects of GK, WK, and article in L1
Table 5.9	Main and interaction effects of grammatical knowledge and world knowledge in advanced L2 speakers
Table 5.10	The log RT residuals by condition in advanced L2
Table 5.11	Main and interaction effects of GK, WK, and article in advanced L2
Table 5.12	Main and interaction effects of grammatical knowledge and world knowledge in intermediate L2
Table 5.13	The log RT residuals by condition in intermediate L2
Table 5.14	Main and interaction effects of GK, WK, and article in intermediate L2

List of Figures

- Figure 2.1 Visual stimulus sample re-produced from Altmann and Kamide (2007)
- Figure 2.2 Visual stimulus sample re-produced from Kamide, Scheepers, and Altmann (2003)
- Figure 2.3 Visual scene re-created from Chambers, Tanenhaus, and Magnuson (2004): Two referent condition
- Figure 2.4 Visual scene re-created from Chambers, Tanenhaus, and Magnuson (2004): One referent condition
- Figure 2.5 Visual scene re-created from Chambers et al. (2002): One referent condition
- Figure 2.6 Visual scene re-created from Chambers et al. (2002): Multiple referent condition
- Figure 2.7 Visual scene sample re-created from Chambers et al. (2002): Second experiment
- Figure 2.8 Visual stimulus re-produced from Trenkic et al. (2013): One-referent condition
- Figure 2.9 Visual stimulus re-produced from Trenkic et al. (2013): Two-referent condition
- Figure 3.1 L1 reading time data before (a) and after (b) logarithmic transformation
- Figure 3.2 Three SD filtered L1 RT distribution before (a) and after (b) log transformation
- Figure 3.3 Log-transformed RT residuals of L1 speakers by condition and region
- Figure 3.4 The interaction of uniqueness and definiteness in the spillover region (L1 data)
- Figure 3.5 Comparison of definite and indefinite NPs by uniqueness (L1 data)
- Figure 3.6 Comparison of unique and non-unique referent conditions by definiteness (L1 data)
- Figure 3.7 Log-transformed RT residuals of advanced L2 speakers by condition and region
- Figure 3.8 The interaction of uniqueness and definiteness in the post-spillover region (L2 data)
- Figure 3.9 Comparison between definite and indefinite NPs by uniqueness
- Figure 3.10 Comparison between unique and non-unique referent conditions by definiteness

Figure 4.1	Visual stimuli re-produced from Ahern and Stevens (2014)
Figure 4.2	Visual stimulus sample for referent prediction task
Figure 4.3	Ratio of unique referent prediction by speaker group and by article
Figure 4.4	The interaction of speaker group and article
Figure 5.1	Visual scene sample re-created from Chambers et al. (2002): Second experiment
Figure 5.2	Visual stimulus re-produced from Trenkic et al. (2013)
Figure 5.3	A visual stimulus example in which world knowledge predicts a unique referent
Figure 5.4	A visual stimulus example in which world knowledge predicts non-unique referents
Figure 5.5	A visual stimulus example for fillers with two plausible options
Figure 5.6	A visual stimulus example for fillers with four plausible options
Figure 5.7	L1 speakers' log reaction time residuals by condition
Figure 5.8	Interaction of grammatical knowledge and world knowledge in L1 speakers
Figure 5.9	L1 speakers' log reaction time residuals by condition and visual stimuli type
Figure 5.10	Advanced L2 speakers log reaction time residuals by condition
Figure 5.11	Interaction of grammatical knowledge and world knowledge in advanced L2 speakers
Figure 5.12	Advanced L2 speakers' log reaction time residuals by condition and visual stimuli type
Figure 5.13	Intermediate L2 speakers' log reaction time residuals by condition
Figure 5.14	Log reaction time residuals by condition and visual stimuli type in intermediate L2
Figure 5.15	Intermediate L2 speakers' log RT residuals with inversed grammaticality
Figure 5.16	Interaction of grammatical knowledge and world knowledge in intermediate L2 speakers

Chapter 1. Introduction

For a long time, researchers have tried to explain why the outcome of second language (L2) learning displays such great variability, unlike first language (L1) learning. One widely held claim involved age of onset for L2 learning (Johnson & Newport, 1989; 1991), which is no longer considered an ‘explanation’ but more of an ‘observation’ that one’s odds of success decrease if one starts late. More recently, research has focused on what makes L2 learning so different and/or difficult. A general consensus that has recently emerged holds that the difficulty of achieving native-like proficiency comes from the need to integrate information from different domains in real time (Clahsen & Felser, 2006; Phillips & Ehrenhofer, 2015; Sorace, 2011; Sorace & Filliaci, 2006). This idea calls for research that looks into the integration of information from multiple sources.

To investigate the issue, I chose to focus on the interaction of the linguistic information conveyed by English articles with the non-linguistic information that comes from real-world knowledge. The two factors were selected because the English article system is one of the most challenging grammatical features for L2 speakers whose L1 does not have a similar system (Ionin, Ko, & Wexler, 2004; Ionin, Zubizarreta, & Maldonado, 2008; Ionin, Zubizarreta, & Philippov, 2009; Garcia Mayo, 2009), while real-world knowledge is a non-linguistic source of information that is known to affect sentence processing (Altmann & Kamide, 1999; 2007; Chambers, Tanenhaus, Eberhard, Filip, & Carlson, 2002; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995).

It is well documented that multiple factors affect L1 sentence processing (Altmann & Kamide, 1999; 2007; Boland, 2005; Chambers et al., 2002; Novick, Thompson-Schill, &

Trueswell, 2008; Sedivy, Tanenhaus, Chambers, & Carlson, 1999; Tanenhaus et al., 1995 among many others). Listeners are guided not only by linguistic information such as syntactic knowledge (Clifton & Frazier, 1989; Ferreira & Clifton, 1986), semantic knowledge relating to thematic roles (Trueswell, Tanenhaus, & Garnsey, 1994), lexical knowledge (Novick et al., 2008), and prosody (Snedeker & Casserly, 2010; Snedeker & Trueswell, 2003), but also by non-linguistic information such as real-world knowledge (Altmann & Kamide, 1999; 2007), visual scenes (Tanenhaus et al., 1995 among many others), and contextual information (Altmann & Kamide, 2009; Sedivy et al., 1999). Given that sentence processing involves not only linguistic but also non-linguistic information, a question arises as to how linguistic and non-linguistic constraints are processed by L2 speakers. It is evident that mature adult L2 learners' capacity to process and understand real-world knowledge, visual scenes, and contextual information does not differ from that of L1 adults. This dissertation focuses on how L2 learners process linguistic information in the presence of this sort of non-linguistic information.

My research attempts to tease apart the processing of linguistic and non-linguistic information in the course of L2 processing and to test the hypothesis from the existing literature that it is difficult to glean information from multiple domains. I use self-paced reading, referent prediction task, an online survey, and referent identification task to examine how Korean learners of English as a second language can use definiteness as uniqueness as an important linguistic cue to process sentences and to predict upcoming linguistic materials and how integrating the linguistic information with non-linguistic information (i.e., world knowledge) increases the burden on the processor.

Chapter 2 reviews relevant literature on L1 and L2 processing and on definiteness. Chapter 3 shows, via self-paced reading, that the pattern of definiteness processing observed in

L1 speakers also appears in L2 speakers, although they unfold more slowly. Chapter 4 (referent prediction task) demonstrates that definite and indefinite English articles can be used by both L1 and advanced L2 speakers as a cue to predict upcoming words. This chapter also considers the development of English articles in L1-Korean L2 speakers of English, which progresses from using indefinite articles for singularity marking at the intermediate level to correctly distinguishing definite and indefinite articles based on definiteness operationalized as unique indefinability at the advanced level.

Chapter 5 presents two experiments: Experiment 3a—a norming study on world knowledge, and Experiment 3b—referent identification task on the integration of linguistic and non-linguistic information. The norming study shows that L1 and L2 speakers display process real-world knowledge in the same way. Experiment 3b—the integration experiment—shows that L1 speakers integrate both linguistic and non-linguistic information incrementally, generating and revising their expectations based on the information that is made available incrementally. In contrast, advanced L2 speakers failed to use linguistic information when they could rely on non-linguistic information to predict what was going to be said next. The findings have theoretical, pedagogical, and methodological implications of importance.

In terms of theory, the findings suggest that processing efficiency in integrating information from different domains can account for non-targetlike behavior in L2 sentence processing. Experiments 1 and 2 show that L2 speakers have target-like knowledge of the (in)definite semantics of English articles and use that knowledge to predict referents yet to be mentioned. Experiment 3a confirms that L1 and L2 speakers share the same world knowledge, thereby excluding the possibility that the performance of the L2 learners in the integration experiment is due to their lack of grammatical knowledge or world knowledge. The fact that L2

speakers, even at an advanced level, could not use linguistic information to revise their initial, world-knowledge-based, expectation suggests that they attempt to streamline sentence processing by initially focusing their attentional resources on more meaningful and easier-to-process information (world knowledge) rather than on relatively new and less automatized information (English articles). This dovetails with the claims that processing is the key to understanding L2 development.

Methodologically, the findings have implications both for the study of L2 English articles and for psycholinguistics research. The current study employs online experimental methods, unlike studies of English articles that rely on off-line methods (Ionin, Ko, & Wexler, 2004; Ionin, Zubizarreta, & Maldonado, 2008 *inter alia*, cf. Trenkic, Mirkovic, Altmann, 2013), whose results often ended up reflecting participants' metalinguistic knowledge. For psycholinguistics, referent prediction task and referent identification task provide simpler alternatives to eye tracking. They are much easier to implement and less prone to methodological issues of eye-tracking (see Chapter 4 for details).

Pedagogically, my results have implications for instructional methods such as content based instruction (Lyster, 2007, Lyster, & Ballinger, 2011; Lyster & Ranta, 1997), task-based language teaching (Long, 2016; Van Den Branden, Bygate, and Norris, 2009) and processing instruction (Van Patten, 2002; Van Patten, Collopy, Price, Borst, and Qualin, 2013). Especially, the findings account for the psycholinguistic mechanisms underlying Processing Instruction as Van Patten and his colleagues also argue for L2 learners' strategy of giving priority to information that makes a greater contribution to interpreting the meaning of a sentence. Further research on the effect of real-world knowledge in L2 pedagogy will shed light on the development of instructional methods and materials.

Chapter 2. Background

This chapter begins with a general overview of second language processing literature that motivates the main research question of the dissertation: How are native language (L1) and second language (L2) behavior different (or similar) in integrating linguistic and non-linguistic information? The chapter then proceeds on to the rationale for why the constructs ‘linguistic information’ and ‘non-linguistic information’ are operationalized by the definiteness distinction underlying the English article system and real-world knowledge, respectively. The discussion of the motivation for the research is followed by a review of L1 and L2 acquisition literature on the English article system, discussion of definitions of definiteness in the semantics literature, and a survey of studies on the role of real-world knowledge in sentence processing. The chapter concludes with research questions to be pursued in the following chapters.

2.1 Investigating processing to explain learning

In recent years, ‘processing’ became the buzzword in both L1 and L2 acquisition research. Phillips and Ehrenhofer (2015) elaborated on how processing research can shed light on learners, learners’ proficiency, and developmental processes, and they called for research that investigates how learners process grammatical contingencies where information should be integrated from multiple linguistic domains. O’Grady (2015) argues that the pressure to minimize processing cost is a driving force for L1 acquisition and explains many phenomena on the developmental trajectory of L2 acquisition. He divides such processing pressures into two types—internal (efficiency-related) factors and external (input-related) factors—and puts forward a

bold claim that all the variation observed with regard to L1 and L2 acquisition can be explained under a uniform powerful force: processing efficiency.

In the L1 psycholinguistics literature, research findings have been reported that parsers process information in an incremental manner to generate and revise expectations in real-time as different sources of information become available (Altmann, 1999; Chambers, Tanenhaus, Eberhard, Filip, & Carlson, 2002; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). Some of them investigated the integration of multiple linguistic domains such as syntax and semantics (Clifton, Traxler, Mohamed, Williams, Morris, & Rayner, 2003; Trueswell, Tanenhaus, & Garnsey, 1994) or syntax and prosody (Snedeker & Trueswell, 2003) while others involved the integration of linguistic and non-linguistic information. Some examples of non-linguistic information will be visual contexts, real-world knowledge, and story contexts. Attachment decisions are affected by visual contexts (Tanenhaus et al., 1995), real-world knowledge circumscribes referential domains (Chambers et al., 2002), and story contexts guide the assignment of thematic roles (Altmann, 1999).

This dissertation focuses on the integration of linguistic and non-linguistic information in L2 sentence processing, not on the integration of different linguistic sources of information, and the rationale for that is as follows. First of all, one can safely assume that mature L2 adults will not differ from mature L1 adult participants in the psycholinguistics literature when it comes to processing non-linguistic information such as visual contexts, real-world knowledge, and story contexts. On the other hand, it has been reported that L2 speakers could not use article information to general expectations (Grüter, Lew-Williams, & Fernald, 2012; Lew-Williams & Fernald, 2010, cf. Hopp, 2016) or they are slower in processing linguistic information (Trenkic,

Mirkovic, & Altmann, 2013¹). When there is a discrepancy in the efficiency with which each source of information can be processed, the incrementality of sentence processing reported in L1 literature will create a different situation for L2 sentence processing.

Linguistic and non-linguistic information are operationalized as the definite semantics of English articles and real-world knowledge in order to maximize the discrepancy in the efficiency with which each type of knowledge is processed. The English article system is one of the most challenging features of English grammar to L2 learners. Even at a highly-advanced level and with an extended period of exposure to natural input, learners experience great difficulty using definite and indefinite articles properly (Garcia Mayo, 2009; Ionin, Ko, & Wexler, 2004; Lardiere, 1998). As mentioned above, L2 adults are not expected to have difficulty incorporating real-world knowledge in sentence processing.

When L2 speakers' real-time processing of linguistic information is expected to be either deficient or inefficient, examining the role of non-linguistic information will shed light on whether 'processing' can explain learning, as claimed in recent literature (O'Grady, 2015; Phillips & Ehrenhofer, 2015; Sorace, 2011).

2.2 Definiteness in L1 and L2 acquisition and processing research

The acquisition of articles in English and other languages is one of the most intensively researched topics in both first language acquisition (Bohnacker, 1997; Demuth, 2001a, 2001b;

¹ Please note that Trenkic et al.'s (2013) experiments were interpreted to show Chinese L2 speakers' ability to anticipate upcoming linguistic material. Indeed, the participants showed anticipatory eye movements in three out of four conditions. However, the illustrations of participants' eye movements clearly show that the L2 speakers' anticipatory eye gazes were delayed compared to L1 speakers.'

Gerken, 1991; 1994; 1996a, 1996b; Gerken & Landau, 1990; Gerken & McIntosh, 1993; Goad & Buckley, 2006; Guasti, Gavarró, de Lange, & Caprin, 2008; Kupisch, Anderssen, Bohnacker, & Snape, 2009; Tremblay, & Demuth, 2007; Wexler, 2011 among many others) and second language acquisition (Blom, Polisenska, & Weerman, 2008; Butler, 2002; Ekiert, 2004, 2010; Goad & White, 2009; Huebner, 1979; 1983; Ionin, 2016; Ionin, Baek, Kim, Ko, & Wexler, 2012; Ionin, Ko, & Wexler, 2004, 2007; Ionin & Montrul, 2010; Ionin & Zubizarreta, 2010; Ionin, Zubizarreta, & Maldonado, 2008; Hopp, 2011; Jarvis, 2002; Kim & Lakshmanan, 2009; Ko, Ionin, & Wexler, 2004, 2010; Liu & Gleason, 2002; Master, 1987, 1997; Murphy, 1997; Parrish, 1987; Robertson, 2000; Sheen, 2007; Snape, Leung, & Ting, 2006; Tarone, 1980, 1983; Tarone & Parrish, 1988; Trenkic, 2004, 2007, 2008; Trenkic & Pongpaiboj, 2013; Zdorenko & Paradis, 2008, 2011).

In the L1 literature, many researchers investigated the role of prosody in the emergence of initial article-like mono syllables (Demuth, 2001a, 2001b; Gerken, 1991, 1994, 1996a, 1996b; Goad & Buckley, 2006; Tremblay & Demuth, 2007) unlike in L2 literature, where the scope of their foci range over prosody (Goad & White, 2009), semantics (Ko et al., 2004, 2010), L1 transfer (Ionin & Montrul, 2010; Ionin et al., 2008; Robertson 2000; Snape et al., 2006), the role of salience (Trenkic & Pongpaiboj, 2013) and the role of task types (Tarone, 1980, 1983; Trenkic, 2007)—just to name a few. In terms of methodology, there have been naturalistic studies that reported longitudinal data on the development of English articles (Huebner, 1979, 1983) to more controlled studies that targeted to measure the development of L2 English article system (Butler, 2002; Ionin et al., 2012; Ionin et al., 2004, 2007; Tarone & Parrish, 1988).

When there are countless previous studies on the topic, one might wonder why it should be looked at again or whether there is anything left to be discussed. Prior to discussing the

construct “definiteness” and the existing literature on L1 and L2 acquisition and processing of definiteness, it is important to discuss why definiteness has been chosen as the focus for this dissertation. There are four reasons.

First, the English article system in L2 processing should be looked at because of its important role in maintaining a coherent referent throughout the discourse between speaker and hearer. An NP headed by the noun *student* can refer to different entities depending on the situation in which the NP is uttered. For example, in both (2.1) and (2.2), a specific student is discussed but in the former, speaker B has no knowledge of the person that *a student* refers to whereas, in the latter, the two interlocutors understand that *the student* refers to the one who didn’t submit his/her assignment.

(2.1) A: You said you have office hours from now till 2pm?

B: Yes, I have *a student* coming in for the 1:30–2:00pm slot.

(2.2) A: Everyone in my class submitted their assignment on time except one.

B: Will you give *the student* an extension?

Secondly, such extremely subtle grammatical marking of definiteness as in the English article system (*the* vs. *a/an*) does not exist in all languages, which makes it one of the most challenging aspects of English grammar for second language learners to master for those who do not have an equivalent system in their L1. Even quite proficient L2 speakers of English rarely reach the native-like mastery of English articles and often suffer from incorrect classroom instructions on English articles (Butler, 2002; Tryzna, 2009). In contrast, L2 speakers of English

whose L1 has a similar article system are reported not to display great difficulty in using English articles correctly (Garcia Mayo, 2009).

Thirdly, a great majority of studies focused on production data, which may not reflect the development of Interlanguage adequately. Studies that tapped into online experimental methods to investigate the processing of definite and indefinite NPs are quite rare in both L1 and L2 but significantly more so in L2. Kim and Lakshmanan (2009) examined whether L2 speakers would behave differently using a self-paced reading task to compare the online data with results from an offline method (acceptability judgment task). Unfortunately, their experimental design was seriously flawed in that (1) the contextual information that determines whether a definite or an indefinite NP should be used came after the critical NP and (2) the reading time data analysis they used dealt with regions that could not shed light on the processing of critical NPs. The current research will be one of very few studies that use online experimental methods to investigate L2 English article use.

Finally, one of the most influential strands of research by Ionin and her colleagues (Ionin et al., 2004, 2007; Ko et al., 2004, 2010; Ionin et al., 2008) used experimental stimuli that confounded specificity with particular types of sentences that could have provided metalinguistic cues to participants. They use the notions of definiteness and specificity and argue that the interference of specificity was pointed to as a culprit of the (relative) failure of L2 speakers' native-like article use. In Chapter 3, I will elaborate on the issues of how the findings reported in their studies need to be re-examined due to their methodological weaknesses, and describe a self-paced reading task that attempts to fine-tune what was previously observed in L2 English article research.

2.3 The semantics of definiteness

Defining definiteness. Defining definiteness has been a long-standing challenge (Hawkins, 1991; Heim, 1982; Russell, 1905; inter alia). The two most widely cited in the discussion of definiteness are Heim (1982) and Russell (1905): The former argued for definiteness as familiarity while the latter advocated definiteness as uniqueness—the view adopted here.

Definiteness as familiarity. Heim (1982) saw mental discourse as a card filing system where mutually known discourse referents between speaker and hearer exist as card files. Each existing discourse referent is represented by a card and when a card from the file is drawn, that is, when a discourse referent is mentioned from the common ground, a definite NP is used. On the other hand, a newly introduced referent should be put on an empty card and included into the card file. When it is first brought into discourse, it will be referred to by an indefinite NP but once it is put into the card file and retrieved for later mentions, it will be referred to by a definite NP. The critical distinction here is whether a referent is (or is assumed to be) familiar to both interlocutors in a conversation. An example can be drawn from Murphy's (1984) study. In (2.3a), both Steve and George were passed by the same truck whereas in (2.3b), the two people were passed by different trucks.

(2.3) Though driving 55, Steve was passed by *a truck*.

a. Later, George was passed by *the truck*, too.

b. Later, George was passed by *a truck*, too.

In the case of (2.3a), the card for the definite NP ‘the truck’ is drawn from the card file. Because ‘a truck’ was introduced in the first sentence, the common ground between interlocutors (writer and reader in this case) is established that there is a truck that is familiar to both of them that exists in the mental discourse of both parties. And the second mention of the referent is given as the definite NP *the truck*. In contrast, in (2.3b), the indefinite NP *a truck* indicates that the entity to which it refers (i.e., the truck that passed George) is not the same truck that passed Steve. The second mention is not familiar to both parties of the conversation and the referent does not exist in the card file between the interlocutors.

Definiteness as uniqueness. Russell (1905) characterized a definite NP in terms of reference to a uniquely identifiable referent. That is, if there is a unique referent that is believed to be identifiable by both speaker and hearer, it will be referred to by a definite NP. For example, when there is mutual understanding that a kitchen is a place where a single, thus, uniquely identifiable, stove is typically expected, using a definite NP is felicitous in referring to the stove (2.4a) (Clifton, 2013). Along the same vein, using an indefinite NP in the same context (2.4b) is infelicitous in that it violates the shared presupposition of a single stove in a typical kitchen. On the other hand, in the context of an appliance store, where multiple stoves are expected, an indefinite NP is felicitous while a definite NP is not (2.5).

(2.4) Singular context: a unique referent presupposed

- a. In the kitchen, Jason checked out *the stove* very carefully.
- b. In the kitchen, Jason checked out *#a stove* very carefully.

(2.5) Multiple context: multiple possible referents presupposed

- a. In the appliance store, Jason checked out *#the stove* very carefully.
- b. In the appliance store, Jason checked out *a stove* very carefully.

The two notions of definiteness have been at the center of controversy for decades, and many researchers present various theoretical grounds to argue which one defines definiteness better. Discussing which captures the notion of definite semantics is beyond the scope of this dissertation. For the current research, definiteness will be operationalized as uniqueness or unique identifiability for a practical reason: “Definiteness as familiarity” could tap into learners’ metalinguistic knowledge. That is, if experimental conditions can be differentiated in terms of the presence versus the absence of repeated NPs,² as in (2.6a) versus (2.6b), that in and of itself can hint at what participants are expected to do. By using the concept of definiteness as uniqueness (2.4–2.5), we can avoid the confound that the repetition of NPs or lack thereof creates in an experiment.

- (2.6)
- a. I bought a car last year. *The car* is used mostly by my husband.
 - b. She walks everywhere. She has *a car* but rarely drives it.

² Please note that this does not indicate that Heim’s definiteness as familiarity concerns only repeated NPs. However, operationalizing definite and indefinite NPs as the presence and absence of previous mentions defines definiteness as interlocutors’ familiarity to a referent.

2.4 The role of real-world knowledge in L1 and L2 sentence processing

2.4.1 Studies on real-world knowledge in L1 sentence processing

For more than a decade, the role of world knowledge (WK) in sentence processing has been actively researched. World knowledge has mostly been looked at for its influence on the processing of formal grammar: phonology (Sedivy, Tanenhaus, Chambers, & Carlson, 1999), morphology (Altmann & Kamide, 2007; Kamide, Scheepers, & Altmann, 2003), syntax (Chambers, Tanenhaus, & Magnuson, 2004) and semantics (Chambers, Tanenhaus, Eberhard, Filip, & Carlson, 2002., Sedivy et al., 1999; Altman & Kamide, 1999; Kamide, Altmann, & Haywood, 2003). In rare cases, world knowledge per se was looked at for its role in sentence processing (Cook & Meyers, 2004).

The research methods varied from visual world paradigm eye-tracking (Altmann & Kamide, 2007; Chambers et al., 2002; Chambers et al., 2004; Kamide, Altmann & Haywood, 2003; Kamide, Scheepers, & Altmann, 2003; Sedivy et al., 1999) to eye-tracking while reading (Cook & Meyers, 2004) and to even related potentials (Hagoort, Hald, Bastiaansen, & Petersson, 2004; Martin, Garcia, Breton, Thierry, & Costa, 2014; Martin, Garcia, Breton, Thierry, & Costa, 2015; Nakano, Saron, & Swaab, 2010). Most of the studies focused on the simultaneous integration of information from linguistic and non-linguistic domains but there were other studies that compared sentences with semantic violation and world knowledge violation in separate trials (Hagoort et al., 2004; Martin et al., 2014, 2015; Nakano et al., 2010).

Most of the studies above aimed to explore the extent to which sentence processing is incremental. For example, studies such as Altmann and Kamide (2007) or Kamide, Scheepers,

and Altmann (2003) looked into how morphological processing interacts with world knowledge. (2.7) is an example of auditory stimuli used in Altmann and Kamide (2007). They were accompanied by a visual stimulus as in Figure 1.

(2.7) The man will drink (*or has drunk*) the beer (*or wine*).



Figure 2.1. Visual stimulus sample re-produced from Altmann and Kamide (2007)

The verb *drink* filters out undrinkable objects such as cheese and crackers. With the two potential referents remaining, the tense morphology on the verbs predicts what will be the complement of the verb — *beer* or *wine*. The results show that participants use real-world knowledge to generate expectations for upcoming linguistic material. A glass becomes empty when someone has drunk its contents and a verb phrase such as *will drink* makes sense when a glass is full. Significantly more anticipatory looks were observed for the empty beer glass when the auditory stimuli had the present perfect form (*has drunk*) than when it was in the future tense (*will drink*) and vice versa for the wine glass case. This study shows that real-world knowledge plays an important role in the processing of linguistic information (tense morphology).

Kamide, Scheepers, and Altmann (2003) used the relatively free word order of German to

investigate the relationship between morphology and real-world knowledge in generating and revising predictions for upcoming linguistic material. They used auditory stimuli as in (2.8) along with a visual stimulus as in Figure 2.2. In (2.8), the two sentences differ in their word orders. In (2.8a), the first NP is the subject of the sentence while, in (2.8b), the first NP is the object. Nominative and accusative case marking can be found both in the definite articles preceding the head nouns and the word-final inflections. *Hase* is a masculine noun that takes *der* for the nominative case and *den* for the accusative case. It also inflects the word ending to *Hasen* in the accusative case.

- (2.8) a. Der Hase frißt gleich den Kohl.
The hare-NOM eats shortly the cabbage-ACC.
“The hare will shortly eat the cabbage.”
- b. Den Hasen frißt gleich der Fuchs.
The hare-ACC eats shortly the fox-NOM.
“The fox will shortly eat the hare.”

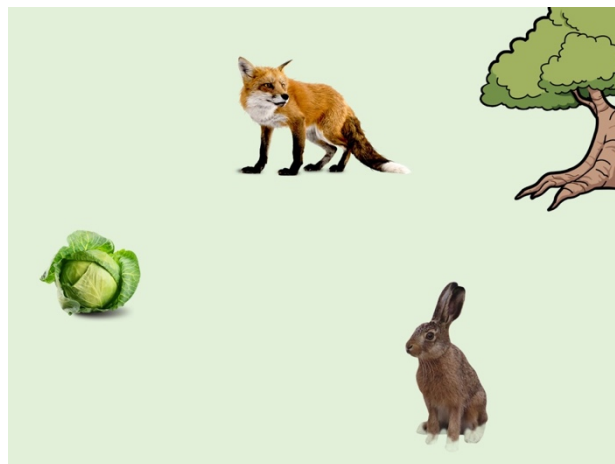


Figure 2.2. Visual stimulus sample re-produced from Kamide, Scheepers, & Altmann. (2003).

When the hare is given as the subject of the verb ‘eat’ (*frißt*) (2.8a), participants’ eye-gazes showed anticipatory movements towards the cabbage. This is similar to the English construction “The rabbit eats...” and real-world knowledge tells us that the object of the verb is more likely the cabbage (*den Kohl*) than the fox (*der Fuchs*). Therefore, participants look to the cabbage even before the actual word is heard.

When the hare was inflected for the object of the verb ‘eat’ (*frißt*) (2.8b), however, anticipatory eye-gazes towards the fox were observed. *Den Hasen frißt* is similar to the passive English construction “The rabbit is eaten.” The accusative case on the NP reveals that the rabbit is the patient of the eating action, which allows participants to predict that the next word of the sentence will be the agent of the action. According to our world-knowledge, a cabbage cannot eat a rabbit but a fox can. This study shows that native speakers use not only linguistic information (case marking) but also real-world knowledge actively and incrementally when generating expectations for the remaining obligatory argument of the verb.

Both Altmann and Kamide (2007) and Kamide, Scheepers, and Altmann (2003) above show how morphological processing interact with real-world knowledge to build anticipations incrementally for linguistic material yet to be presented. Chambers et al. (2004) shows how syntactic parsing decisions are also affected by world knowledge. Auditory stimuli as in (2.9) were counterbalanced with two visual stimuli (Figures 2.3 and 2.4). The difference between the two visual stimuli was whether one or both of the eggs were in the liquid form. In one of the visual stimuli, both containers had egg in the liquid form (Figure 2.3). In the other case, only one egg was in the liquid form and the other was still in the shell so it could not be poured (Figure 2.4). The contrast created the two-referent and one-referent conditions.

- (2.9) a. Pour the egg in the bowl over the flour.
b. Pour the egg that's in the bowl over the flour.

The prepositional phrase (PP) 'in the bowl' in (2.9a) is temporarily ambiguous between the goal interpretation (i.e., into the bowl) and the modifier interpretation (that is in the bowl) whereas, the relative clause *that's in the bowl* in (2.9b) prevents such ambiguity. Facing the ambiguity as in (2.9a), the goal interpretation is preferred; however, the interpretation of the prepositional phrase was affected by whether both eggs were pourable or only one of them was (i.e., whether the visual stimulus included one or two pourable eggs).

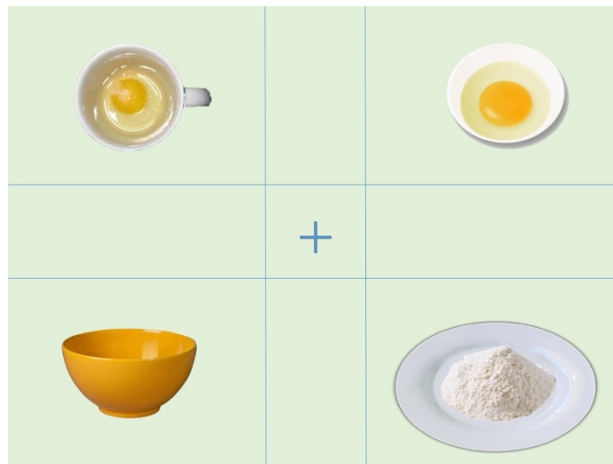


Figure 2.3. Visual scene re-created from Chambers et al. (2004): two referent condition.



Figure 2.4. Visual scene re-created from Chambers et al. (2004): one referent condition

When both were in the liquid form (Figure 2.3), participants interpreted the PP as a modifier to single out one referent of the two. In contrast, when only one of the eggs was in liquid form (Figure 2.4), participants thought the thematic role of the PP was the *goal*. That is, real-world knowledge informs participants that they don't need a modifier to pick out which egg should be poured because there is only one pourable egg. In sum, world knowledge plays an important role in syntactic parsing as well.

Chambers et al. (2002) calls for special attention in this section. They report two separate experiments that tested the role of real-world knowledge in sentence processing and the second experiment directly deals with processing definite and indefinite NPs. Their first experiment looked at how lexical items such as *inside* and *below* could manipulate the range of potential referents.

(2.10) Pick up the whistle and hold it over the cross. Now put it below/inside the can.

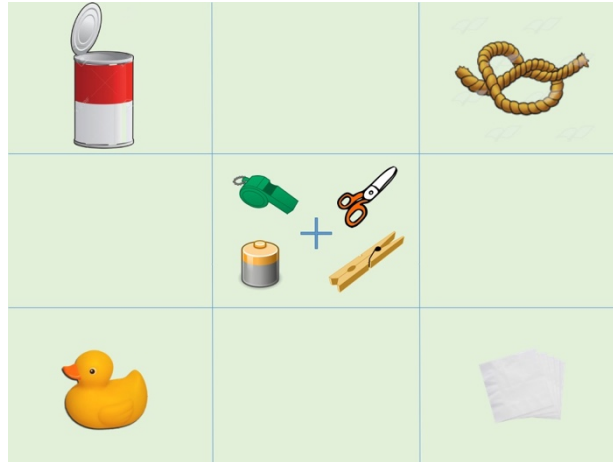


Figure 2.5. Visual scene re-created from Chamber et al. (2002); One-referent condition



Figure 2.6. Visual scene re-created from Chamber et al. (2002); Multiple-referent condition

Auditory stimuli as in (2.10) were accompanied by visual stimuli where there were small theme objects in the center of the 3 x 3 grid, and four larger goal objects in the four corner cells of the grid (Figures 2.5 and 2.6). Visual stimuli had two different conditions. One of them had only one container-like goal object: a can (Figure 2.5). The other stimulus had three container-like goal objects: a bowl, a cup, and a can (Figure 2.6). This way, they could examine the role of world knowledge in processing the prepositions *below* and *inside*. The preposition *below* suited

all four goal objects but *inside* sounds natural only when goal objects are container-like.

Participants fixated their eye gazes on the target faster when there was only one container-like goal object (Figure 2.5) than when there were three container-like objects (Figure 2.6). They concluded that participants use their world knowledge to generate expectations for what is coming up next in the sentence.

Their second experiment looks at the interaction of (in)definite NPs and world knowledge. For auditory stimuli, the second experiment used only *inside* but used two different articles for the goal objects: *the* vs. *a* (2.11). Visual stimuli manipulated the goal objects in a different way from the first experiment. There were a total of six objects placed alongside the perimeter of a circle (Figure 2.7): a theme (a cube), three goal objects (a small can, a large can, and a bowl) and two distractors (a rubber duck and a hammer).

(2.11) Put the cube inside the/a can.

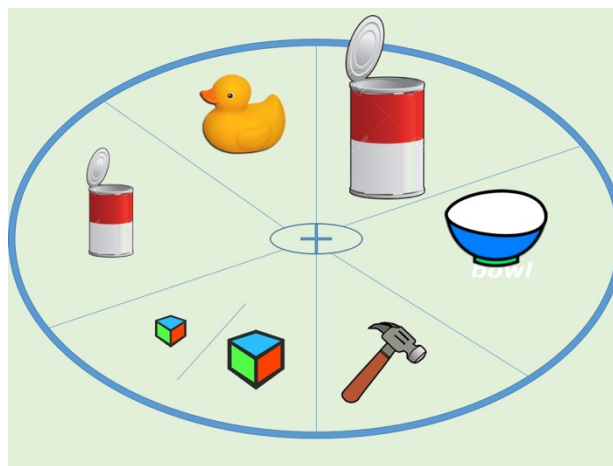


Figure 2.7. Visual scene sample re-created from Chambers et al. (2002); Second experiment

The important difference between the first and second experiments was that the second

experiment manipulated the number of goal objects by using theme objects of different sizes. In one condition, the theme object (cube) was small and could fit in both small and large cans (two-referent condition) whereas, in the other, the theme object was large and could fit only the large can (one referent condition). When the world knowledge dictated that the large cube would fit only one of the cans, the definite NP *the can* led to earlier fixations to the target (one referent condition). The indefinite NP *a can*, however, induced earlier fixations to the target when the small cube could fit both cans (two referent condition). That is, the size of the theme object and the size of the goal object were considered to narrow down the potential candidates for the critical NP (*the can* vs. *a can*). The results indicate that native speakers use their WK in processing linguistic information.

The studies surveyed so far all explored how real-world knowledge could influence the interpretation of linguistic information (morphology, syntax, and semantics) and both linguistic and non-linguistic information were presented within one trial. There are a few studies that compared semantic processing and world-knowledge processing in separate trials using ERP (Haggort et al., 2004; Martin et al., 2014, 2015; Nakano et al., 2010). An example from Martin et al. (2014) illustrates how semantic and world knowledge were compared. Whereas (2.12a) is a fully acceptable sentence, (2.12b) involves a semantic anomaly and (2.12c) a WK error.

- (2.12) Before the age of eight, children start to ... and write.
- | | | |
|----|-------|---------------------------|
| a. | read | <i>correct</i> |
| b. | bark | <i>semantic violation</i> |
| c. | smoke | <i>WK violation</i> |

They found out that (1) semantic violations induced significantly smaller P2s or the positive deflection of waveforms of ERP peaking around 100–250 ms after the stimulus (Sur & Sinha, 2009) than correct and WK-violation sentences and that, (2) as for N400 or the negative deflection peaking around 300–600 ms post-stimulus (Sur & Sinha, 2009), all three conditions were different from each other with correct sentences eliciting the smallest N400 amplitude and semantic violations the greatest. The results suggest that native speakers respond more sensitively and earlier to a semantic violation than to a WK violation.

2.4.2 Studies on real-world knowledge in L2 sentence processing

Only a very limited number of studies have looked at the role of WK in L2 sentence processing using online methods such as eye-tracking and ERP (Mitsugi & MacWhinney, 2015; Trenkic, Mirkovic, & Altmann, 2013). Here, I will review in detail Trenkic et al. (2013) that replicate Chambers et al. (2002) and elaborate on why the study is not an ideal design for the purpose of the current dissertation: investigating the integration of information from multiple sources.

Trenkic et al. (2013) replicated the second experiment of Chambers et al. (2002). They collected data from intermediate-level Mandarin Chinese speakers using pictures on a computer screen instead of actual objects unlike in the original study. Another difference was that, instead of the size of the theme object (a small cube or a large cube), the goal object was manipulated as open and closed to differentiate the number of potential targets (Figures 2.8 and 2.9).

In Figure 2.8, one container is open but the other container is closed and in Figure 2.9, both containers are open. This difference in the visual scene is related to world knowledge that

one cannot put an object inside a closed can. When processing auditory stimuli as in (2.13), participants need to use their world knowledge and decide how many containers are potential candidates for the goal argument.

(2.13) The man will put the cube inside the/a can.

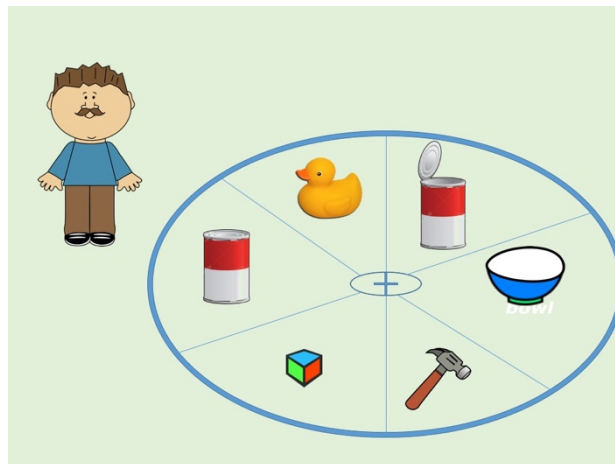


Figure 2.8. Visual stimulus re-produced from Trenkic et al. (2013); One-referent condition.

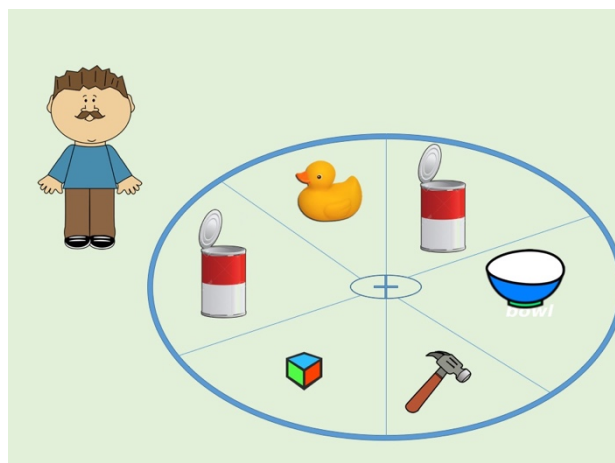


Figure 2.9. Visual stimulus re-produced from Trenkic et al. (2013); Two-referent condition.

Trenkic et al. (2013) replicated the results of Chambers et al. (2002) with L1 speakers: L1 speakers can fixate on the target faster when the uniqueness of referents matched the definiteness of the NPs. That is, the definite NP *the can* led to earlier fixations on the target in the one-referent condition (Figure 2.8) and the indefinite NP *a can* led to earlier fixations in the two-referent condition (Figure 2.9). L2 speakers were slower than L1 speakers in general but they also showed the same pattern of mapping unique referents to definite NPs and non-unique referents to indefinite NPs. Trenkic et al. (2013) concluded that L2 speakers could use real-world knowledge in processing definite and indefinite NPs in a native-like manner and discuss the results in comparison with previous studies where participants varied greatly in the production of English articles.

All the studies reviewed above have important implications in psycholinguistics: they show that sentence processing does not rely solely on linguistic information and that both linguistic and non-linguistic information is actively involved. However, the studies above do not tease apart the role of linguistic information and world knowledge. It is clear that world knowledge played a role in sentence processing in both L1 and L2; however, the studies did not include cases where linguistic information and non-linguistic information clashed. For example, in the second experiment of Chambers et al. (2002) and the only experiment in Trenkic et al. (2013), WK dictates the number of plausible referents; therefore, the use of either a definite or an indefinite article is intertwined with WK. In the one-referent condition (Figure 2.8), *the* is grammatical and *a* is ungrammatical and vice versa in the two-reference condition (Figure 2.9). However, this is only a violation of grammatical knowledge but not of the world knowledge. All the other studies of world knowledge reviewed here show that WK plays a role in sentence processing but they do not examine the effect of information clash. This design therefore cannot

be used here given that the aim of this dissertation is to compare the role of grammatical knowledge and that of world knowledge.

Therefore, I devised a novel experiment design in which linguistic information and non-linguistic information are pitted against each other. In the last experiment of this study (to be elaborated on in Chapter 5), linguistic information indicates ‘definiteness as uniqueness’ manifested in the English article system and non-linguistic information is represented by ‘world knowledge (WK)’ that dictates (for example) that a scientist is more closely related to a microscope than to a sweater and a little child will prefer a lollipop over a bell pepper. The main goal of this study is to show how the two are integrated. Prior to conducting the integration experiment, however, we must determine whether both L1 and L2 speakers have the same linguistic and non-linguistic information. Therefore, studies that look into grammatical knowledge (English articles) and world knowledge will be independently conducted before the integration experiment.

2.5 Research Questions

The first two of four experiments will delve into L1 and L2 processing of definiteness as uniqueness; to show that L2 speakers share the same linguistic knowledge as L1 speakers. Experiment 1 in Chapter 3 uses a self-paced reading task for which experimental design and items were borrowed and modified from Clifton (2013). This experiment will test how L1 and L2 speakers can presuppose the indication of unique or non-unique referents and accommodate definite and indefinite NPs, respectively. The experiment will shed light on participants’ knowledge on definiteness as unique identifiability. Experiment 2 in Chapter 4 will inverse the

order of presentation and see if either a definite or indefinite article will lead to the prediction of a unique or a non-unique referent. The third experiment is an online survey to confirm that both L1 and L2 speakers of English have the same type of WK. The type of real-world knowledge used in this experiment is relationships between people (or animals) and objects. Associations between a doctor and a stethoscope (as opposed to a laptop) and between a scientist and a microscope (instead of a sweater) will be used to operationalize the construct ‘real-world knowledge.’ The last experiment will investigate how the two different sources of information are integrated in real-time(Experiment 3b in Chapter 5).

RQ1. Does definiteness indicate uniqueness in L1 sentence processing? Do L2 speakers have the same grammatical knowledge of definiteness as uniqueness as L1 speakers?

(Chapter 3: Experiment 1)

RQ2. Do L1 and L2 speakers use definite and indefinite articles to predict a unique referent?

(Chapter 4: Experiment 2)

RQ3. Do L1 and L2 speakers share the same type of World Knowledge (WK)?

(Chapter 5: Experiment 3a)

RQ4. When linguistic information and non-linguistic information indicates a different referent (i.e., a definite NP indicates a unique referent but WK indicates a non-unique referent or vice versa), which information will prevail in L1 sentence processing? Will L2 speakers show the same prevalence?

(Chapter 5: Experiment 3b)

Chapter 3. Processing of definiteness as uniqueness

3.1. Introduction

The purpose of this chapter is two-fold. One is to confirm that, prior to investigating the integration of linguistic and non-linguistic information, both first language (L1) and second language (L2) speakers have access to the same linguistic information in sentence processing. The other is to re-visit the claim that universal semantics plays a role in L2 processing of (in)definiteness and examine the validity of the studies that presented such a claim. For these purposes, the current chapter investigates L1 and L2 processing of ‘definiteness as uniqueness’ in linguistic information. The focus is on how L1 and L2 speakers react to the mapping between the definiteness of a noun phrase (NP) and the uniqueness of a referent.

In L1 research on definiteness processing, different claims have been made with regard to the processing cost of definite and indefinite NPs. Murphy (1984) found that indefinite NPs are costlier to process than its definite counterparts. Schumacher’s studies (Burkhardt, 2006; 2008; Schumacher³, 2009) focused on the role of presupposition in processing definite NPs and argued that not all definite NPs are equal in terms of the degree to which their interpretations are dependent upon discourse contexts. Clifton (2013) argued that definite and indefinite NPs do not lead to significantly different processing cost when unique and non-unique referents are presupposed for definite and indefinite NPs, respectively.

In L2 research on English article use, the role of specificity was one of the most commonly posited factors in accounts for non-target-like L2 behavior (the Fluctuation

³ Please note that Burkhardt and Schumacher refer to the same person.

Hypothesis; Ionin, Ko, & Wexler, 2003; 2004). However, the strong argument has subsided as Samoan, which Ionin and colleagues used as evidence of a natural language that marks specificity via their article system, turned out not to make such a distinction. More recently, different results have been documented between online and offline experiments (Ionin, Baek, Kim, Ko, & Wexler, 2012; Ionin, Kim, & Tyndall, in progress; Ionin, Zubizarreta, & Phillippov, 2009). Yet, such discrepancy between online and offline tasks could be better explained as the interference of incorrect metalinguistic knowledge that had been reported in Butler (2002).

This chapter replicates the findings from Clifton's self-paced reading experiment. (In)definiteness and (non-)uniqueness interact not only in L1 English but also in L2 English; hence, definite and indefinite NPs do not differ in terms of processing cost as long as previous contexts presuppose unique and non-unique referents, respectively. Also, pairwise comparisons between matching and mismatching conditions suggest that both L1 and L2 speakers are sensitive to misused definite NPs but not to errors regarding indefinite NPs. The only difference between L1 and L2 speakers is that the interaction effect and the unbalanced pairwise comparison results were delayed in L2 speakers in comparison to L1 speakers. The results suggest that the fluctuation effect observed in previous studies could be an artifact of their experimental design.

3.2. Background

3.2.1. L1 Research on definiteness processing

This section reviews three strands of psycholinguistic research that are most relevant to the main goal of the dissertation and that looked into the online sentence processing of definiteness in L1 speakers. The survey leads to motivations for the current study.

Murphy (1984) reported a reading time difference between definite and indefinite NPs, and the mental discourse model was employed to account for the observed reading time discrepancy. (3.1) shows how the definite NP *the truck* (3.1b) and the indefinite NP *a truck* (3.1c) are distinguished as given and new under the context of (3.1a). Murphy argues that processing a definite NP indicates accessing an existing referent in the mental discourse whereas processing an indefinite NP means introducing a new entity into the mental discourse; hence, the reading time difference.

- (3.1) a. Context: Though driving 55, Steve was passed by a truck.
 b. Given: Later, George was passed by *the truck*, too.
 c. New: Later, George was passed by *a truck*, too.

Murphy's experiments focused on the givenness of an NP. In (3.1b), *the truck* refers to the same truck as that in the context sentence (3.1a). The indefinite NP *a truck* in (3.1c), on the other hand, refers to a truck that is different from the one previously mentioned. Murphy does not address the case in which both definite and indefinite NPs could be new. In the right contexts, however, a definite article can accompany a noun that is first introduced in the discourse.

- (3.2) I took a taxi this morning and *the driver* was very kind.

In (3.2), the definite NP *the driver* is mentioned for the first time in the sentence but one can infer that there must be a single, unique driver in any given taxi; thus, using the definite article is well accommodated by a parser. When a non-unique entity, which could be any one of many referents, is presupposed, an indefinite NP will be used. In (3.3), anyone with an appropriate driver's license can be the driver for the said bus; hence, an indefinite NP is used.

(3.3) The bus needs someone to drive it. Let's hire *a driver*.

The definite NP *the driver* in (3.2) and the indefinite NP *a driver* in (3.3) are both new to the discourse, so their difference lies in the presupposition of a uniquely identifiable referent, not in the givenness of an entity. From this observation, the question arises whether definite and indefinite NPs will still differ in terms of processing cost when the context provides a cue on the (non-)uniqueness of a referent. A prerequisite to answering that question is to determine whether presupposition does play a role in definiteness processing.

Schumacher's studies (Burkhardt, 2006; Burkhardt, 2008a; 2008b; Schumacher 2009) provide an answer to a part of that question. (3.4) is one of her examples (Burkhardt, 2006) where the first three sentences (3.4a–3.4c) contextualize the target sentence (3.4d) differently. In the identity condition, the NPs *a conductor* in (3.4a) and *the conductor* in (3.4d) are identical; in the inference condition, the context sentence (3.4b) presupposes a uniquely identifiable conductor; and lastly, in the incoherent condition, the context sentence (3.4c) provides no clue for the definite NP *the conductor* in the target sentence (3.4d).

- (3.4) a. Identity: Tobias visited *a conductor* in Berlin.
- b. Inference: Tobias visited *a concert* in Berlin.
- c. Incoherence: Tobias talked to Nina.
- d. Target: He said that *the conductor* was very impressive.

The results of Schumacher's Event Related Potential (ERP) experiments indicate that the presupposition of a referent plays a role in online processing. A definite NP like *the conductor* has a significantly pronounced N400⁴ when it is used randomly with no contextual presupposition (3.4c) as opposed to the two previous conditions that includes an identical referent *the conductor* (3.4a) or presupposes a uniquely identifiable referent: *a concert* (3.4b)⁵. Her findings present strong evidence that "presupposition" about a referent makes an important difference in accommodating a newly introduced definite NP.

In addition to Burkhardt (2006), her later studies (Burkhardt, 2008a; 2008b) also examined the extent to which a referent is dependent upon discourse contexts; however, they did not discuss indefinite NPs and it was only Schumacher (2009) that included an indefinite condition as in (3.4e)

- (3.4) e. Indefinite: He said that *a conductor* was very impressive.

⁴ Pronounced N400 indicates a subject's detection of semantic incongruity (Sur & Sinha, 2009).

⁵ A comparison between the identity condition (3.3a) and the inference condition (3.3b) bridges a gap between Murphy (1984) and Clifton (2013). A definite NP that is directly identical to the previous mention (Murphy, 1984) and a newly introduced but contextually inferable definite NP (Clifton, 2013) do show a discrepancy in processing cost. A significantly more pronounced P600 was observed in the latter. This is not discussed in detail as it is not closely relevant to the current study.

She reports a main effect of context in N400, which was also observed in all of her previous studies and an interaction effect of context and definiteness for late positivity. Pairwise comparisons on the late positivity⁶ effect showed that a significant difference was observed only between definite-identity and indefinite-identity conditions and in no other pairs of conditions. The study did not clarify if presupposing a context where there is a unique or a non-unique referent can help accommodate definite and indefinite NPs. Therefore, the question remains unanswered as to whether the difference in processing cost between definite and indefinite NPs could be compensated for by contextual presuppositions.

Clifton (2013) tackles this question and reports that definite and indefinite NPs per se are not significantly different in terms of reading time as long as contexts introduce a presupposition about the (non-)uniqueness of a referent. Examples in (3.5) and (3.6), borrowed from Clifton's study, illustrate how (non-)unique referents can match and/or mismatch (in)definite NPs.

(3.5) In the kitchen, (a uniquely identifiable stove is presupposed)

- a... Jason checked out *the stove* very carefully. Definite NP (match)
- b... Jason checked out *a stove* very carefully. Indefinite NP (mismatch)

(3.6) In the appliance store, (no uniquely identifiable stove is presupposed)

- a... Jason checked out *the stove* very carefully. Definite NP (mismatch)
- b... Jason checked out *a stove* very carefully. Indefinite NP (match)

⁶ Late positivity is related to processing cost that arises from updating the discourse.

A kitchen is a place where a single (thus, unique) stove is typically expected (3.5) whereas an appliance store typically has multiple (thus, non-unique) stoves (3.5). The contexts are followed by either a definite or an indefinite NP, which creates four different conditions. In the unique referent condition, a definite NP matches the context (3.5a); in the non-unique condition, an indefinite NP matches the context (3.6b). An indefinite NP in a unique referent context (3.5b) and a definite NP in a non-unique referent context create mismatch conditions (3.6a).

Clifton's experiments on self-paced reading and eye-tracking while reading uncovered a significant interaction between uniqueness and definiteness. This means that, when used in the matching contexts, definite and indefinite NPs do not result in any significant reading time difference. The results, however, do not discuss pairwise comparisons between matching and mismatching conditions: That is, his report does not provide the statistics for the direct comparisons between individual conditions, e.g., a comparison between (3.5a) and (3.5b), between (3.6a) and (3.6b), between (3.5a) and (3.6a), and between (3.5b) and (3.6b). A significant difference in such comparisons will provide stronger evidence that definite NPs and indefinite NPs are in a one-to-one mapping relationship with unique referents and non-unique referents, respectively.

The main quest with regard to describing L1 online processing of definiteness will be to examine if definite and indefinite NPs elicit significantly different (or the same) processing cost when unique and non-unique referents are presupposed. The two research questions (RQs) are,

- (3.7) RQ1. Does the presupposition about the (non-)uniqueness of a referent and the (in)definiteness of a referring NP interact?

RQ2. Will pairwise comparisons between matching and mismatching conditions show a significant difference in terms of processing cost in both unique and non-unique referent contexts?

3.2.2. Research on L2 English article use

Amongst the vast array of research on L2 English article use, this chapter zooms in onto a series of studies by Ionin and her colleagues, which were some of the most influential in L2 English article research in recent years (Ionin et al., 2012; Ionin et al., in progress; Ionin et al., 2004; Ionin et al., 2009). Their earlier studies put forward the Fluctuation Hypothesis (FH) (Ionin et al., 2003; 2004 inter alia) and in later observations, they discovered L2ers' non-target-like behavior or fluctuation is observed only in offline tasks that exploit explicit knowledge (Ionin et al., 2009). More recently, they are investigating whether L1 transfer can account for L2 English article use. In this section, I will provide an in-depth examination of their arguments and suggest how the apparent role of specificity observed in their earlier studies could have resulted from the way 'specificity' was operationalized in their experimental items.

Ionin et al.'s (2004) study used the semantic notions of definiteness and specificity and argued for the role of specificity in L2ers' fluctuating performance in applying definiteness to distinguish definite and indefinite NPs. The sentences in (3.8) are examples from Lyons (1999) that illustrate the specificity distinction of indefinite NPs.

- (3.8) a. Peter intends to marry *a merchant banker* though he doesn't get on at all with her.
b. Peter intends to marry *a merchant banker* though he hasn't met one yet.

In (3.8a), there is a specific woman that Peter has in mind for his future wife [+specific]. On the other hand, no specific female merchant banker is in Peter's life yet in (3.8b); any member of the set "female merchant bankers" satisfies a prerequisite Peter has for his spouse yet unspecified [–specific]. This contrast illustrates how the same indefinite NP *a merchant banker* could either refer to a certain entity or have no concrete referent. Examples in (3.9) are a modified version of experimental items from Ionin et al. (2004).

- (3.9) a. Now the race is over, I'd like to interview the winner; look how excited she is!
 b. When the race is over, I'd like to interview the winner; who do you think will win?

The same specificity contrast also exists in definite NPs. There is a certain person who won the race in (3.9a) [+specific]. But, in (3.9b), the game is still on going and no specific person has been set to be the winner yet [–specific]. As can be seen from the two examples, there is no grammatical marker to distinguish specificity in English.

The English article system distinguishes only definiteness (Table 3.1). According to Ionin et al. (2004), Samoan was previously known to use its two articles *le* and *se* to mark specificity (Table 3.2) and that L2ers of English fluctuate between the two parameters of the universal semantic distinction (definiteness vs specificity) (Table 3.3). Ionin et al. (2004) reported that L2 speakers' article choice fluctuated when the feature values of definiteness and specificity clashed ([+definite, –specific] or [–definite, +specific]) (the shaded cells of Table 3.3). The observed fluctuation, however, needs reconsidering because of the way they manipulated the [±specific]

feature. (3.10) to (3.13) are examples of experimental items used in Ionin's forced choice elicitation task.

Table 3.1. *Definiteness distinction in English*

	+def	–def
+spec	<i>the</i>	<i>a/an</i>
–spec		

Table 3.2. *Specificity distinction in Samoan*

	+def	–def
+spec	<i>le</i>	
–spec	<i>se</i>	

Table 3.3. *Fluctuation between definiteness and specificity in Interlanguage grammar*

	+def	–def
+spec	+	+/–
–spec	+/–	–

Participants were asked to choose one of the three options in the parentheses before the target NPs. The correct answer for each condition is *the* for the [+definite] conditions—(3.10) and (3.11); and *a* for the [–definite] conditions—(3.12) and (3.13).

(3.10) [+definite, +specific] 12 items

Police Officer C: I haven't seen you in a long time. You must be very busy.

Police Officer S: Yes. Did you hear about Miss Sarah Andrews, a famous lawyer who was murdered several weeks ago? We are trying to find (a, *the*, —) murderer of Miss Andrews—his name is Roger Williams, and he is a well-known criminal.

(3.11) [+definite, –specific] 8 items

Reporter: Several days ago, Mr. James Paterson, a famous politician, was murdered! Are you investigating his murder?

Police Officer: Yes. We are trying to find (a, *the*, —) murderer of Mr. Peterson—but we still don't know who he is.

(3.12) [–definite, +specific] 8 items

Reporter1: Hi! I haven't seen you in weeks. Do you have time for lunch?

Reporter 2: Sorry, no. I'm busy with a story about local medicine. Today, I am interviewing (*a*, the, —) doctor from Bright Star Children's Hospital—he is a very famous pediatrician, and he doesn't have much time for interviews. So I should run!

(3.13) [–definite, –specific] 12 items

Professor Clark: I'm looking for Professor Anne Peterson.

Secretary: I'm afraid she is busy. She has office hours right now.

Professor Clark: What is she doing?

Secretary: She is meeting with (*a*, the, —) student, but I don't know who it is.

A careful look at the items, however, reveals that specificity was systematically marked by other means. [+specific] was indicated by providing ample information about target NPs and [-specific] was hinted on by denying any knowledge of target NPs. Sixteen out of the twenty [+specific] items provided either the name of the referent denoted by the target NP or included a sentence such as “She is a friend of mine”; (3.10) and (3.12). Also, twelve out of the twenty [-specific] items included a variety of “I don’t know the person”; (3.11) and (3.13).

This sharp contrast between the two specificity conditions might as well have misled participants when they selected answers. This was exacerbated by the facts (1) that the task was in a paper-and-pencil test format, which allowed participants to go back and forth between items, (2) that there were no fillers, and (3) that all the cues followed, not preceded, the target NPs, which made the results even more irrelevant to real-time language use⁷. The issues with the experiment materials, the task type, and the presentation order of target NPs and their contextual cues weaken the claim of the Fluctuation Hypothesis (FH) that L2 speakers alternate between the two parameter settings. Such experimental flaws, however, do not completely disprove the role of specificity. Rather, they motivate the need to improve research method to examine if this intriguing phenomenon truly exists.

More compelling evidence against the FH comes from Ionin et al. (2009). They report that a more recent inquiry on the Samoan article system presents a different article distinction from the earlier version they used. The clear-cut specificity distinction demonstrated in Table 3.2 was not seen in the new version (Table 3.4). Their 2009 study used the same materials as in from

⁷ To process target NPs, participants had to read materials that come after them. Such order of presentation is expected to lead readers to tap into their metalinguistic knowledge.

their earlier studies (2003, 2004) to test L1-Russian L2ers of English and found the same fluctuation pattern.

Table 3.4. *Revised le/se distinction in Samoan.*

	+def	–def
+spec	<i>le</i>	
–spec		<i>se</i>

Since Samoan can no longer provide evidence for the universal semantic distinction and the two parameter settings, they discuss a new observation of different results between online and offline experiments: The said fluctuation was observed mostly in offline tasks. Since naturalistic production (online) data lack such fluctuation, they argue that participants' tapping into explicit knowledge could be related to the phenomenon. However, this only highlights the suspicion that the lowered performance in the [+definite, –specific] and [–definite, +specific] conditions could be due to the flawed manipulation of experimental items that misled participants⁸.

Although the FH failed to provide an adequate explanation for non-target-like L2 behavior, the overuse of *the* in [–definite, +specific] NPs has not been observed only in Ionin's studies. The phenomenon is one of the most-documented in adult L2 research (Ahn, 2009; Butler, 2002; Huebner

⁸ In fact, Butler (2002) documents that L1-Japanese L2ers of English listed incorrect strategies in deciding which article to use. E.g., When there is a modifier, use *the*. Such misleading strategies are commonly discussed (and even encouraged) in English test prep classes of East Asian countries.

1979) as well as in L1 and L2 child language development. What sets apart Ionin's earlier findings (Ionin et al., 2003; 2004) from all other research on the same topic is that none of the latter studies reported the overuse of *a(n)* in [+definite, –specific] NPs. If the overuse of *a(n)* can be observed in a well-designed laboratory experiment, specificity could be an adequate explanation for errors in L2 English article use.

Ionin's focus shifted from the Fluctuation Hypothesis to L1 transfer in Ionin et al. (2012) in which L1-Korean participants are reported to use *the* when a Korean demonstrative *ce*⁹ /tʃʌ/ (a Korean equivalent of English *that*) is needed. This so-called cross-linguistic influence, however, is also observed only in offline tasks that tap into explicit knowledge. In online tasks, L2ers behave more like L1 speakers (Ionin et al., in progress). If their performance of using Interlanguage grammar were influenced by their L1 grammar, it would more likely to appear in online tasks when they tap into implicit knowledge.

In brief, both arguments (the role of specificity and L1 transfer) need further investigation in a more scientifically rigorous manner. The current study will shed light on the first issue using the same self-paced reading method as in Clifton (2013). The two research questions from the L1 section above (3.7) will provide guidelines to investigate the issue. The first question is about the interaction of (non-)uniqueness and (in)definiteness. As long as they are in matching pairs ([+unique, +definite] and [–unique, –definite]), definite and indefinite NPs should not display a significant reading time difference. The interaction effect in L2 speakers will indicate that definite and indefinite NPs are at least distinguished in L2 grammar.

The second question is about pairwise comparisons between matching and mismatching conditions in each context. As can be seen in (3.5) and (3.6), target NPs in all experimental items

⁹ Yale romanization is used for transliteration.

of the current study are [–specific]. According to the Fluctuation Hypothesis, in a [–specific] context, more errors are made with definite articles than with indefinite articles (see Table 3.3). An error in a production task translates into less sensitivity to an error in a receptive task (i.e., self-paced reading). If the fluctuation phenomenon in Ionin et al. (2004) is not an outcome of a methodological oversight but is a key factor in L2 English article use, participants should show less sensitivity to errors in definite articles and more sensitivity to those in indefinite articles. In statistical terms, pairwise comparisons should show a significant difference between unique and non-unique contexts in indefinite conditions, but not in definite conditions.

3.3. Method

3.3.1. Self-paced reading (SPR)

Self-paced reading experiments measure reading times (RTs) of each word or phrase in a sentence to probe readers' reaction to grammatical anomalies. The method was developed in the 1970s by several psycholinguists (Aronson & Scarborough, 1976; Mitchell & Green, 1978) and was first introduced to SLA research by Juffs and Harrington (1995). A sentence is presented either word by word or phrase by phrase, using what is often called the “moving window technique”—so-called because it limits the focus of the eye gaze to a small number of words at a time. The resulting eye movement is similar to when we are looking out a window, which allows us to see only a portion of the view outside. As participants press a button, the window moves to the next word (or phrase) and a computer program measures the time that a reader stays in each window. The region where a target grammatical anomaly is embedded is called the critical

region, and RTs usually increase one region after a grammatical anomaly is detected by the reader. The region after the critical region is called the “spillover” region. Because each participant presses the button to move on to the next window at his or her own pace, it often happens that the reader does not react immediately to the anomaly. Thus, the RTs of the critical region are not as much affected as the RTs of the spillover region, where the surprise takes effect.

- (3.14) a. Cumulative design
- | | |
|------------------------------|-------------------------------|
| At the beginning of a trial: | --- ----- |
| At the first button press: | The ----- |
| At the second button press: | The rabbit --- ----- |
| At the third button press: | The rabbit ran ---- ----- |
| At the fourth button press: | The rabbit ran into --- ----- |
| At the fifth button press: | The rabbit ran into the ----- |
- b. Non-cumulative design
- | | |
|------------------------------|-------------------------------|
| At the beginning of a trial: | --- ----- |
| At the first button press: | The ----- |
| At the second button press: | --- rabbit --- ----- |
| At the third button press: | --- ----- ran ---- ----- |
| At the fourth button press: | --- ----- --- into --- ----- |
| At the fifth button press: | --- ----- --- ----- the ----- |

Self-paced reading can be either cumulative (3.14a) or non-cumulative (3.14b).

Typically, the letters making up the forthcoming words in a sentence will be shown as dashes. At

the first button press, the first word will be revealed with the remaining words still shown as dashes. The difference between cumulative and non-cumulative self-paced reading is whether, at the next button press, the previous word will remain shown on the screen or disappear and be replaced with dashes again. The current experiment used the non-cumulative design to ensure that RTs in each region did not include time it takes for eye gazes to go back to previous words and linger there.

3.3.2. Proficiency

A C-test (Schultz, 2006) was used to measure L2ers' proficiency but was also administered to L1ers for a comparison purpose. The test was made up of two paragraphs, each of which had 20 blanks. The blanks were placed on the latter half of every other content words. Both L1ers and L2ers were given 15 minutes maximum to work on the C-test; the statistics of each group's C-test scores are summarized in Table 3.1.

3.3.3. Participants

A total of 134 participants participated in the experiment. Fifty-four L1 English speakers participated as a native control group, and 80 L1-Korean L2-English speakers were recruited in Honolulu, Hawai'i in the U.S. and Seoul, Korea. Course credit or monetary compensation was provided. Table 3.5 summarizes the demographics of the participants. The L2 participants were divided into advanced and intermediate groups by their C-test scores. The native language of L2 participants was limited to Korean. Because Korean does not have an article system, unlike

German or Spanish, good performance on the task by the Korean L2ers cannot be attributed to their L1 knowledge.

Table 3.5. *Participant demographics*

		Age (years)	C-test (out of 40)	LOR (years)
L1 (n=55)	<i>Mean</i>	25.96	35	N/A
	<i>SD</i>	9.26	2.3	
	<i>Range</i>	18–57	28–39	
Advanced L2 (n=38)	<i>Mean</i>	26.22	32.21	3.86
	<i>SD</i>	6.95	2.5	4.14
	<i>Range</i>	20–41	29–38	0–6
Intermediate L2 (n=42)	<i>Mean</i>	25.51	25.5	0.09
	<i>SD</i>	4.15	2.4	0.21
	<i>Range</i>	19–32	21–28	0–0.75

3.3.4. Stimuli

Critical items. Experiment stimuli were created by modifying items from Clifton’s (2013) study and adding more items. Some of Clifton’s items required cultural knowledge that may not be shared by L2ers (e.g., “Staples” may not be recognized as a stationery store by L1-Korean L2ers of English whose length of residence in the U.S. was too brief). Some items were replaced altogether and others were modified. To increase the number of observations, new items were created in the same pattern as Clifton’s for a total of 20 experimental items.

Each item occurred in two different contexts (unique vs. non-unique) and was counterbalanced with definite and indefinite articles. Four lists were created for a Latin-square design. That is, each list included 20 items total, with five items in each of the four conditions. No item was repeated in different conditions within the same list. Items were divided into sections, or regions, as shown in (3.15). Each participant’s reading time at each region was

measured by the experimental software E-prime 2.0 (Psychology Software Tools, Pittsburgh, PA). The critical region was region 6 (R6) in all items.

	R1/R2 /R3	/R4 / R5	/R6	/R7 / R8
(3.15) a.	In /the /kitchen,	/Jason / checked out	/ <i>the stove</i>	/ very / carefully.
b.	In /the /kitchen,	/Jason / checked out	/ <i>a stove</i>	/ very / carefully.
c.	In /the /appliance store,	/Jason / checked out	/ <i>the stove</i>	/ very / carefully.
d.	In /the /appliance store,	/Jason / checked out	/ <i>a stove</i>	/ very / carefully.

Fillers. Fifty-two filler items were included (3.16) (see Appendix 1C). Their sentence structures were similar to those of the critical items, but the critical region did not include the (in)definite article contrast.

(3.16) After finishing lunch, Melody read a novel and took a nap until James came home.

3.3.5. Procedure

Participants were surveyed on their language background through an online questionnaire prior to coming to the lab. In the lab, they were seated in sound-attenuated booths. Instructions and items were presented on a computer screen. Each item was presented using the moving-window technique. Once an entire sentence was presented, a simple arithmetic problem as in (3.17a) followed (as in Clifton, 2013). Participants were told to use the number keypad to answer the question. This simple math question was followed by a comprehension question (3.9b) to

check whether participants had read the sentence carefully. The insertion of an arithmetic problem between a sentence and its comprehension question follows Clifton's (2013) procedure. Neither the self-paced reading task nor the eye-tracking experiment in Clifton's study had significant effects when there was no other task between experimental sentences and their corresponding comprehension questions.

(3.17) a. $34+3=?$

b. Question: What was Jason checking out?

(a) Something he could cook with

(b) Something he could clean with

When the same experiments were run again with arithmetic problems inserted between sentences and comprehension questions, significant differences were observed between predicted pairs of conditions. Clifton's account for the different results is that the added burden on working memory in the experiments with arithmetic problems made participants read more carefully, which, in turn, resulted in more careful attention to such subtle morphological cues as articles. In other words, readers have to read a sentence more carefully if they are going to keep it in memory until they finish a math problem and see the comprehension question. Because the critical regions (*the stove* vs. *a stove*) in the current study's materials are almost the same in length in all four conditions, and articles are short function words that can be easily ignored, it seemed reasonable to use such a technique to ensure careful reading. An arithmetic question

occurred between the test sentences and the comprehension questions in all experimental items and fillers.

3.4. Analysis

3.4.1. Mixed effects modeling

For statistical analysis, linear mixed effects regression (lmer) was used. The benefits of this method is that it can take into account factors such as individual differences in experiment participants and items in one model. Unlike the traditional repeated measure Analysis of Variance (ANOVA), which requires two separate F values for a subject analysis (F_1) and an item analysis (F_2), an lmer model will include both fixed effects and random variances in one model. According to Clark (1973), the two separate analyses of F_1 and F_2 in ANOVA, should be accompanied by F' to strengthen the reliability of analysis. A significant F_1 value indicates that the same results will be replicated when new subjects are recruited given the same items. That is, F_1 shows that the results are not due to the particular set of participants in the experiment. Likewise, F_2 shows that the same results will be replicated when different language materials are used given the same participants. It means that the results do not come from the peculiar set of items used in the experiment. Lastly, F' is the analysis for the case in which both subjects and items are replaced. Most psycholinguistic research that uses ANOVA leaves out this third value. Having to add this third analysis complicates the process of statistical analysis and it still incorporates only a part of subject and item effects: ANOVA models cannot incorporate both fixed and random effects at the same time addressing both the intercepts and slopes for subjects

and items. Therefore, linear mixed effects modeling was used and all statistical analyses were run in R (R Core Team, 2016) using the lme4 package (Bates, Maechler, Bolker, & Walker, 2015).

3.4.2. Assumptions of linear regression

The analysis of self-paced reading is based on how much time is spent on each region given the length of the region. With reading time change per unit length, coefficients are calculated per experimental condition using linear regression. Among some assumptions made in linear regression, two are very important: the linearity of dependent variables (DV) and independent variables (IV), and normal distribution of errors (Crawley, 2007). In the case of self-paced reading, the relationship between reading time (RT) and word length are dependent and independent variables, respectively. If their relationship is linear, reading time increases as words become longer: The longer the word, the longer the reading time. The other assumption is the normality of errors. *Errors* is another name of residuals, that is, discrepancies between expected and observed values.

Below I elaborate on the two assumptions; self-paced reading (SPR) experiments almost always violate these two important assumptions, but the problem can be resolved through logarithmic transformation and residual reading time analysis.

As in many other natural phenomena, the DV (reading time) and the IV (word length) in SPR are not in a linear relationship. People do not take equal amounts of reading time per character. That is, a person who takes 400ms to read a four-character word “read” will not take

700ms to read “reading.” It is safe to assume that SPR data will almost always violate the linearity assumption.

To discuss the other issue: the normality of errors, one should first provide the definition of errors. Here, I provide an example that can make readers easily understand “errors” or “residuals” in statistical analyses. In a hypothetical reading time regression model, one can come up with a simple linear regression model as in (3.18).

$$(3.18) \text{ Reading Time} = 100ms * \text{number of characters}$$

The regression model predicts that it should take 400ms to read “ball” and 500ms for “balls.” However, the expected value 500ms may not be the same as the observed value of 419ms. Then, the regression model needs to address the difference as in (3.19),

$$(3.19) \text{ Reading Time} = 100ms * \text{number of characters} + e$$

where e indicates the error term. This error or residual should be normally distributed, which means that there should be no more or less variation of errors whether an individual reads slowly or fast. The normality of error terms is affected by two factors: (a) the distribution of DV and IV and (b) the linearity of the relationship between the two. If either DV or IV is severely skewed or their relationship is significantly non-linear, error distribution cannot be normal. As seen in Figure 3.1a, the reading time data of the current experiment is severely skewed and as mentioned above, reading time and word lengths are in a non-linear relationship. Therefore, the normal

distribution of errors cannot be expected in the data. In other words, data in SPR experiments are impacted by these two issues.

To fix the problem of linearity violation and to improve the normality of error distribution, raw RT data are put through logarithmic transformation. Figures 3.1a and 3.1b show the normality of raw reading time data and log-transformed reading time data, respectively. Not only is it visually evident that the RT data is more normally distributed in Figure 3.1b, the skewness and kurtosis values from raw RT have changed to the normal range after the log transformation (Table 3.6).

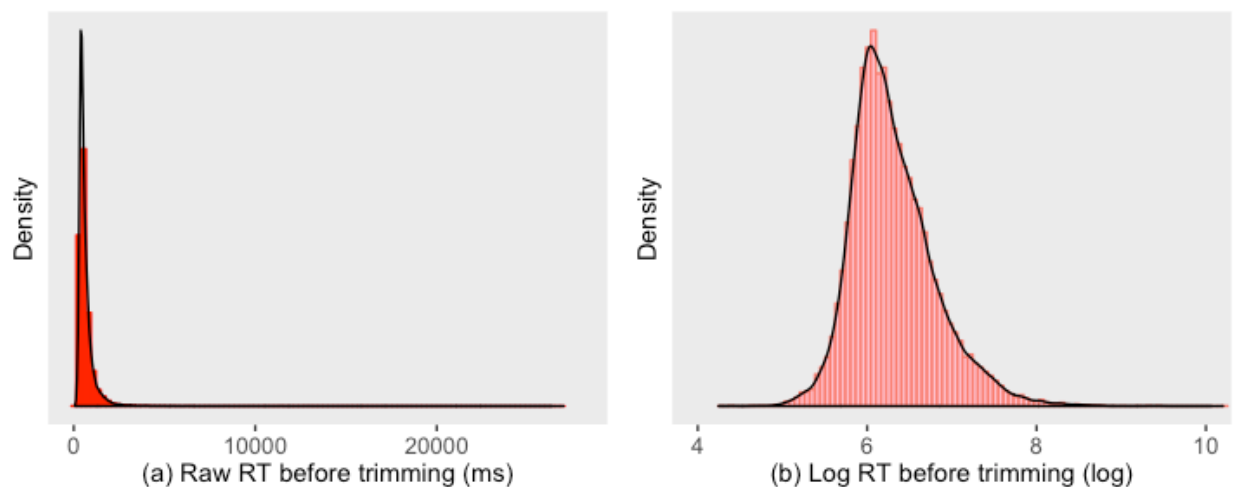


Figure 3.1 L1 reading time data before (a) and after (b) logarithmic transformation.

As can be seen from the x -axis of Figure 3.1a, some of the reading time data was abnormal in that none of the experiment stimuli would have taken as long as 20,000 ms (20 seconds). These data points were outliers that needed to be filtered out. Data points that deviate from the mean of log RT by three times the standard deviation ($3 SD$) or more were filtered out. This resulted in the loss of 0.97% of the L1 data and 0.77% of the L2 data.

Table 3.6. Descriptive statistics of reading time data before and after log transformation.

	L1		L2	
	Raw RTs	Log RTs	Raw RTs	Log RTs
mean	634.17	6.31	742.88	6.46
SD	481.29	0.49	527.51	0.51
Range	70–26983	4.25–10.2	22–10019	3.09–9.21
Skewness	10.29 ($se=0.03^{10}$)	0.87	4.06 ($se=0.02$)	0.81
Kurtosis	350.87 ($se=0.03$)	1.63	33.26 ($se=0.04$)	0.95

Figure 3.2 shows the distribution of raw RT and log RT data after data points were filtered by 3 *SD*. The x-axis of the raw RT data (Figure 3.2b) expands only up to 2500*ms*, which is a reasonable amount of time to read one region in a sentence that consists of one to four words. Even when filtered, the raw RT data is still significantly skewed and logarithms are used throughout the chapter for data analysis.

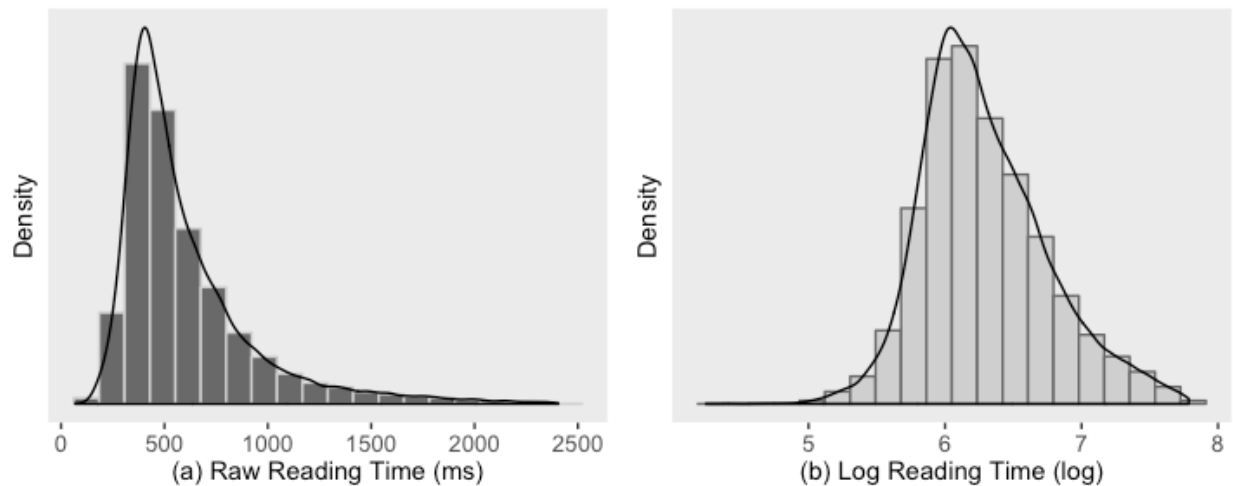


Figure 3.2. Three *SD* filtered L1 RT distribution before (a) and after (b) log transformation.

¹⁰ The *se* values for skewness and kurtosis were calculated following the method provided in Brown (1997).

3.4.3. Residual Analysis

The normality issue is resolved by logarithmic transformation of raw RTs but there are other factors remaining that can affect regression analysis. For example, word lengths vary across conditions, everyone reads at a different pace, and participants speed up towards the end of their session. The current trend in psycholinguistics for solving these issues is to use residual reading time analysis. As was mentioned earlier, residuals are discrepancies between expected and observed values. The residual reading time analysis creates an lmer model that factors into word lengths (*wlen*), order in which items are presented (*order*), and different intercepts for each individual (*subj*) (3.20).

```
(3.20) lmer(logRT~wlen+order+(1|subj), data)
```

Table 3.7. *Linear Mixed Effects Regression output for log RT residual analysis in L1 data*

	Estimate	Std. Error	t value	p value	
(Intercept)	6.1194	0.0324	188.7132	0.0000	***
wlen	0.0390	0.0005	78.7071	0.0000	***
order	-0.0042	0.0001	-43.5877	0.0000	***

· $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$

Table 3.8. *Linear Mixed Effects Regression output for log RT residual analysis in adv. L2 data*

	β (Estimate)	Std. Error	t value	p value	
(Intercept)	6.0560	0.0330	183.7291	0.0000	***
wlen	0.0593	0.0007	81.1451	0.0000	***
order	-0.0030	0.0001	-20.6644	0.0000	***

· $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$

Tables 3.7 and 3.8 show that word length and trial order were affecting reading time significantly in both L1 and L2 data. The model (3.20) used treatment coding in which comparisons are made in reference to a default condition (Intercept). That is, the intercept on the first row indicates the logarithm of the reading time for the shortest word in the data, RTs (log) are estimated to increase when word lengths (*wlen*) become longer and decrease as they become familiar with the task towards the end of the experiment (*order*). With word length and trial order adjusted by participant, residuals are calculated by comparing the expected values from the regression model and the observed values in the actual data.

The resulting residual values are put in mixed effects models to see how they change by experimental condition. The model was run in the *lmer* package using R (3.21a) and its mathematical equation is given in (3.21b). For both L1 and L2 data, maximal *lmer* models that include fixed effects (*definite* and *unique*) and random intercepts and slopes for both participants (*subj*) and experimental items (*item*).

```
(3.21a) lmer(logRTresidual~1+definite*unique+(1+definite*unique|subj)
          +(1+definite*unique|item), data)
```

$$(3.21b) Y_{si} = \beta_{00} + \gamma_{0s} + \delta_{0i} + (\beta_1 + \gamma_{1s} + \delta_{1i}) * X + (\beta_2 + \gamma_{2s} + \delta_{2i}) * W + (\beta_3 + \gamma_{3s} + \delta_{3i}) * XW + e_{si}^{11}$$

Y_{si} is the log-transformed reaction time residual for subject s in item i ;

X is the fixed factor ‘article (*definite*)’;

¹¹ In the model for advanced L2 speakers, the maximal model as in (3.21a) didn’t merge. Subject and item random slopes were included only for the context (*unique*); $Y_{si} = \beta_{00} + \gamma_{0s} + \delta_{0i} + \beta_1 * X + (\beta_2 + \gamma_{2s} + \delta_{2i}) * W + \beta_3 * XW + e_{si}$.

W	is the fixed factor ‘context (unique)’;
β_{00}	is the grand mean of the log-transformed reaction time residual;
γ_{0s}	is the subject (subj) intercept, the deviation from β_{00} for subject s ;
δ_{0i}	is the item (item) intercept, the deviation from β_{00} for item i ;
β_1	is the main effect of the fixed factor X ;
γ_{1s}	is the subject (subj) slope with regard to the fixed factor X ;
δ_{1i}	is the item (item) slope with regard to the fixed factor X ;
β_2	is the main effect of the fixed factor W ;
γ_{2s}	is the subject (subj) slope with regard to the fixed factor W ;
δ_{2i}	is the item (item) slope with regard to the fixed factor W ;
β_3	is the interaction effect of the fixed factors XW ;
γ_{3s}	is the subject (subj) slope with regard to the interaction of XW ;
δ_{3i}	is the item (item) slope with regard to the interaction of XW .

3.5. Results

L1-English speakers showed the interaction pattern in the spillover region, as was predicted based on the previous findings from Schumacher (Burckhardt, 2006; 2008; Schumacher, 2009) and Clifton (2013). Advanced L2 English speakers showed the same interaction pattern as L1 speakers, if delayed one region. As for pairwise comparisons, only one of the two mismatch conditions elicited a significant RT increase in both groups. Intermediate L2 English speakers did not show any significant main effects or interaction effects in any region; hence, the rest of the chapter discusses only the results from L1 and advanced L2 speakers.

3.5.1. L1 speakers

Figure 3.3 shows the log-transformed reading time (RT) residuals of L1 data in the pre-critical region (*checked out*), the critical region (*the/a stove*), the spillover region (*very*), and the post-spillover (*carefully*) region. Residual reading time indicates how much longer it usually takes to read a region with the mean reading time for a region of the same length set as 0. So, negative values indicate a shorter than average reading time, and positive values a longer than average reading time. As the figure indicates, only minimal, insignificant differences are observed in the critical region (*the/a stove*). As was predicted, the effects of interaction between definiteness and uniqueness show up in the spillover region (*very*), where participants took a longer time to process the mismatch conditions (two middle bars) than the match conditions (two bars on the outer edges).

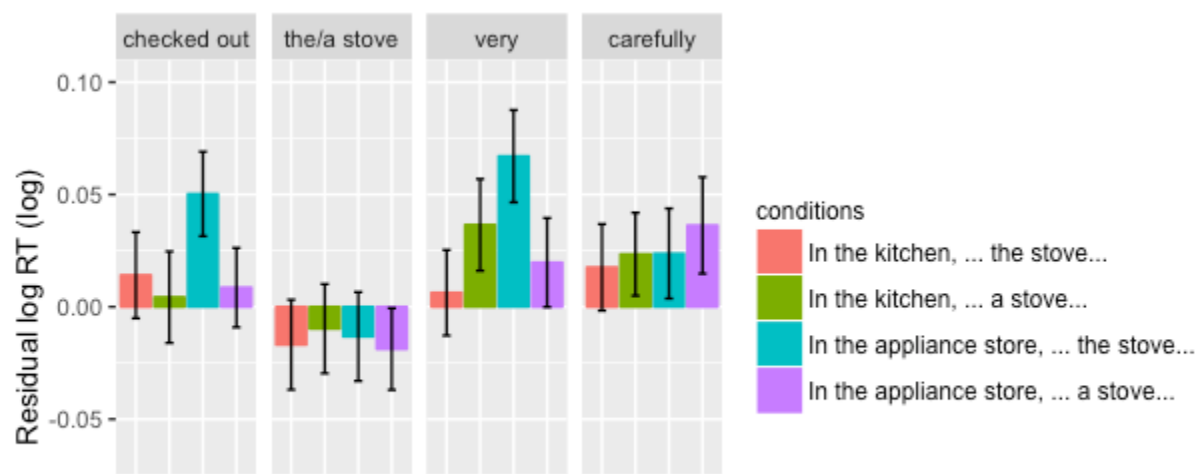


Figure 3.3. Log-transformed RT residuals of L1ers by region.

Note. The error bars indicate standard error.

Table 3.9 reports the slope estimates and their corresponding t values and p values. The analysis used sum coding, which shows the extent to which each categorical variable deviates from the grand mean. The (Intercept) in the first row indicates the grand mean of log RT residuals of the spillover region and the next two rows show the main effects of definiteness (definite) and uniqueness (unique), which are not statistically significant. The last row shows the interaction effect of the two factors, which is significant at .05 alpha level.

Table 3.9. *The lmer output for the interaction between article and context in L1ers.*

(a) Fixed effects

	(Estimate)	Std. Error	t value	p value	
(Intercept)	β_{00}	0.0329	0.0214	1.5346	0.1249
definite (article);X	β_1	0.0036	0.0121	0.2972	0.7663
unique (context);W	β_2	-0.0105	0.0110	-0.9484	0.3429
definite x unique;XW	β_3	-0.0196	0.0095	-2.0567	0.0397 *

. $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$

(b) Random effects

Groups	Name	Variance		Std.Dev.	Corr		
subj	(Intercept)	γ_{00}	0.0046	0.0676			
	definite	γ_1	0.0004	0.0187	0.97		
	unique	γ_2	0.0017	0.0416	-0.65	-0.80	
	definite x unique	γ_3	0.0002	0.0132	0.42	0.61	-0.96
item	(Intercept)	δ_{00}	0.0058	0.0760			
	definite	δ_1	0.0011	0.0326	-0.98		
	unique	δ_2	0.0001	0.0078	-0.04	0.22	
	definite x unique	δ_3	0.0000	0.0026	-1.00	0.99	0.10
Residual			0.0917	0.3027			

Number of obs: 1054, groups: subj, 55; item, 20

The estimates in the second column of the table shows the degree to which each variable contributes to the increase and decrease of log RT residuals. The intercept in Table 3.9 is 0.0329.

This means that the log RT of the spillover region is overall longer than the mean log RT of all regions, which is 6.31 (log) for L1 speakers, as was reported in Table 3.6. The p value for the intercept indicates that it is not statistically different from 0, which, in turn, means that the spillover region does not differ from all other regions in terms of mean reading time.

The estimate for definiteness (*definite*) predicts that log RT residuals will deviate by 0.0036 (log) from the grand mean ($6.31+0.0329+0.0036$ (log)) when a definite NP is given as opposed to an indefinite NP. With an indefinite NP given, log RT will deviate by the same amount only in the negative direction ($6.31+0.0329-0.0036$ (log)).

Likewise, the third row (*unique*) indicates that log RT residuals will decrease or deviate in the negative direction by 0.0105 (log) compared to the grand mean in the unique referent context ($6.31+0.0329-0.0105$ (log)) but increase or deviate in the positive direction by the same amount in the non-unique referent context ($6.31+0.0329+0.0105$ (log)). However, as mentioned above, both article and context did not result in any significant effects.

The negative value of the interaction estimate on the last row indicates that log RT residuals will decrease when unique and non-unique referents are matched with definite and indefinite NPs, respectively. This means that it takes a shorter amount of time to read conditions where articles and contexts match ($6.31+0.0329-0.0196$) compared to when they mismatch ($6.31+0.0329+0.0196$). The graphical illustration that the two outer bars (match conditions) are shorter than the two middle bars (mismatch conditions) helps understand the interpretation of estimated coefficients.

The interaction of definiteness (*definite*) and uniqueness (*unique*), however, should be interpreted with one caveat. As indicated in Figure 3.4, pairwise comparisons between conditions show that not all mismatch conditions are equal in terms of processing cost. The (in)definiteness

distinction (*the* vs. *a*) shows a meaningful difference only in the non-unique referent context (appliance store) and that the (non-)uniqueness distinction (*kitchen* vs. *appliance store*) has a meaningful effect only when it comes to definite NPs (*the stove*). As Figure 3.4 shows, the [–unique, +definite] condition (appliance store, the stove) shows a significant reading time increase compared to the [+unique, +definite] condition (kitchen, the stove) and the [–unique, –definite] condition (appliance store, a stove) ($\beta=0.034$, $SE=0.016$, $t\text{-value}=2.072$, $p=0.038$).

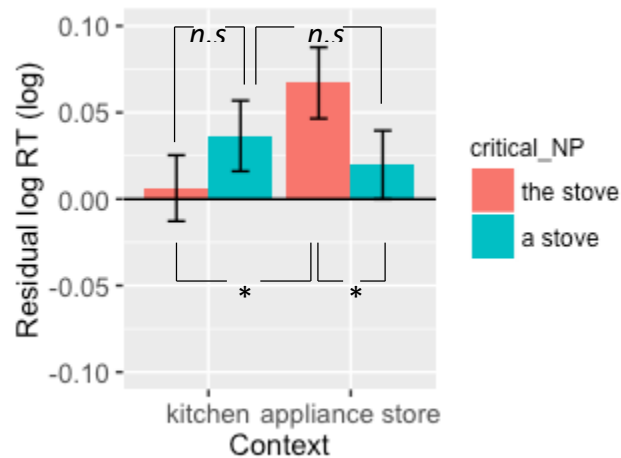


Figure 3.4. The interaction of uniqueness and definiteness in the spillover region (L1 data).

Note. The error bars indicate standard error.

However, the [+unique, –definite] condition (kitchen, a stove) does not elicit as great a reading time increase either from the [+unique, +definite] condition (kitchen, the stove) or the [–unique, –definite] condition (appliance store, a stove) as the [–unique, +definite] condition does ($\beta=0.005$, $SE=0.017$, $t\text{-value}=0.274$, $p=0.784$). Put differently, this means “In the kitchen, Jason checked out a stove very carefully” is not as anomalous as “In the appliance store, Jason checked out the stove very carefully.”

A visual illustration of such discrepancy is provided in Figure 3.5. The pink line indicates estimated log RT residual change of definite NPs between the unique (kitchen) and non-unique (appliance store) conditions and the blue line is for indefinite NPs. The two black dots indicate the mean of log RT residuals in the unique and non-unique contexts. The blue line (indefinite) does not show as great a movement as the pink line (definite) in response to the contexts. As was mentioned earlier, the difference between the definite and indefinite NPs was not statistically significant in the *kitchen* context but significant in the *appliance store* context. Figure 3.6 clearly illustrates how unique and non-unique contexts rarely make a difference for indefinite NPs but make a clear difference for definite NPs.

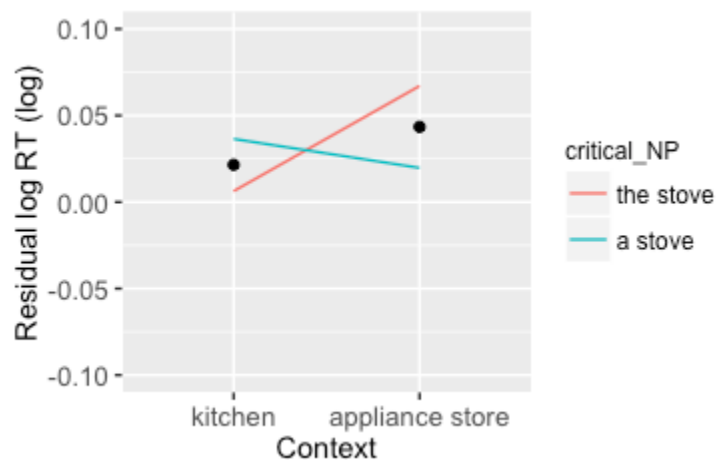


Figure 3.5. Comparison of the definite and indefinite NPs by uniqueness

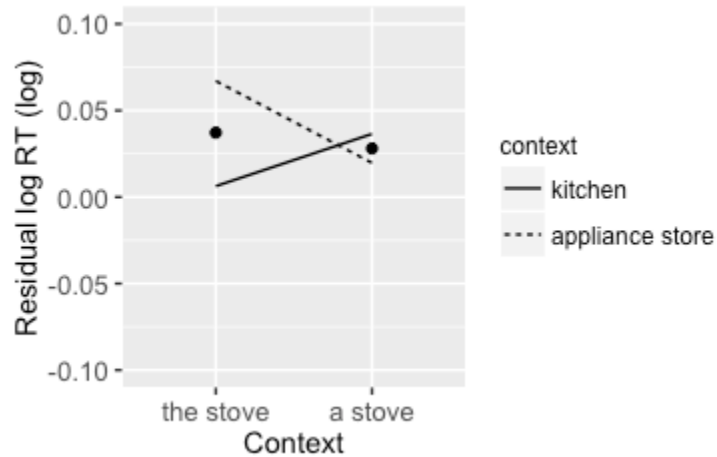


Figure 3.6. Comparison of the unique and non-unique referent contexts by definiteness

The results from L1 data answer the two research questions (3.7). First, the presupposition of a (non-)unique referent and the (in)definiteness of a referring NP do interact. This means definite and indefinite NPs per se do not differ in terms of processing cost when unique and non-unique referents are presupposed, respectively; that is, the two matching conditions are equal in terms of processing cost. Secondly, pairwise comparisons of the L1 data do not show as great a sensitivity to indefinite NPs as they do to definite NPs. The implications of such imbalance will be discussed in 3.6. Discussion.

3.5.2. Advanced L2 speakers

Figure 3.7 shows the log RT residuals of advanced L2 speakers for the four regions from the pre-critical region (“checked out”) to the post-spillover region (“carefully”). Unlike L1ers, the interaction of definiteness and uniqueness was not observed in the spillover region in L2 data; however, the effect was observed one region later in the post-spillover region, which

suggests that advanced L2 speakers react more slowly to grammatical anomalies or infelicitousness. Such delayed reaction, in turn, could result from slower processing in general. As in L1, L2 data showed no significant main effects of either definiteness or uniqueness.

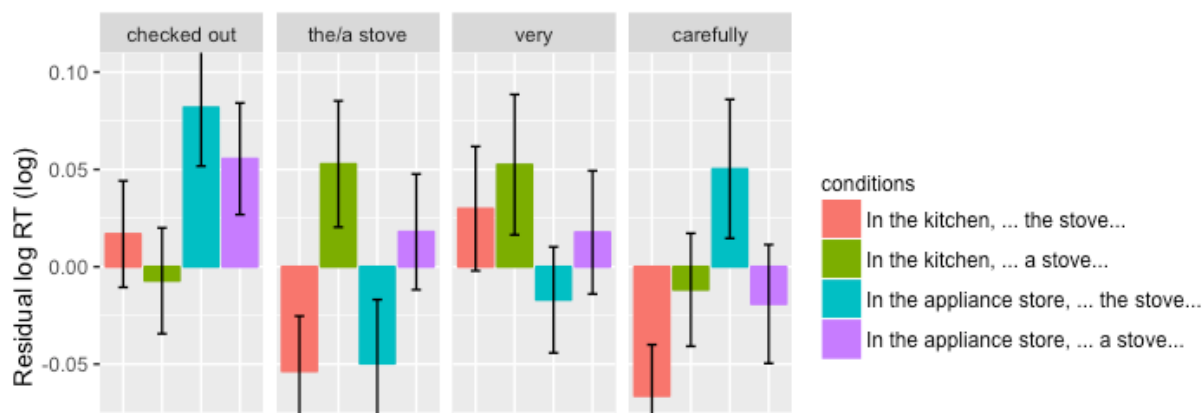


Figure 3.7. Log-transformed RT residuals of advanced L2ers.

Note. The error bars indicate standard error.

Table 3.10 summarizes the lack of main effects (definite and unique) and the significant interaction effect of definiteness and uniqueness in L2 speakers' post-spillover region. This shows that advanced L2 speakers' behavior, if delayed, is similar to that of L1 speakers. Reading the lmer output table is the same as in L1 data. The intercept on the first row of Table 3.10 is – 0.0115. This means that the log RT of the post-spillover region is overall slightly shorter than the mean log RT of all regions, which is 6.46 (log) for advanced L2ers, as was reported in Table 3.2. The p value for the intercept indicates that it is not statistically different from 0, which, in turn, means the post-spillover region itself does not differ from all other regions in terms of mean RT.

The estimate for definiteness (definite) predicts that log RT residuals will deviate by 0.0062 (log) from the grand mean ($6.46 - 0.0115 + 0.0062$ (log)) when a definite NP is given as

opposed to an indefinite NP. With an indefinite NP, log RT will deviate by the same amount only in the negative direction (6.46–0.0115–0.0062 (log)).

Table 3.10. *The lmer output for the interaction between article and context in advanced L2ers.*

(a) Fixed effects

	(Estimate)		Std. Error	t value	p value
(Intercept)	β_{00}	-0.0115	0.0323	-0.3568	0.7212
definite	β_1	0.0062	0.0142	0.4331	0.6649
unique	β_2	-0.0285	0.0164	-1.7389	0.0821
definite x unique	β_3	-0.0285	0.0143	-2.0013	0.0454 *

$p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$

(b) Random effects

Groups	Name	Variance ¹²	Std.Dev.	Corr
subject	(Intercept)	γ_{00}	0.0064	0.0800
	unique	γ_2	0.0057	0.0752 -0.64
item	(Intercept)	δ_{00}	0.0100	0.1000
	unique	δ_2	0.0015	0.0384 1
Residual		0.1146	0.3385	

Number of obs: 578, groups: subject, 30; item, 20

Likewise, the third row (unique) indicates that log RT residuals will decrease or deviate in the negative direction by 0.0285 (log) compared to the grand mean in the unique referent context (6.46–0.0115–0.0285 (log)) but increase or deviate in the positive direction by the same amount in the non-unique referent context (6.46–0.0115+0.0285 (log)). However, as mentioned above, the explanation is only to help readers understand how to read the table. Neither definiteness (definite) nor uniqueness (unique) resulted in any significant effects.

¹² As was explained in foot note #11, the maximal model of random intercepts and slopes did not merge for the advanced L2 speaker data. Hence, random slopes are included only for the context (unique).

The negative value of the interaction estimate on the last row indicates that log RT residuals will decrease when unique and non-unique referents are matched with definite and indefinite NPs, respectively. This means that it takes a shorter amount of time to read matching conditions; [+definite, +unique], [–definite, –unique] ($6.46 - 0.0115 - 0.0285$) compared to when they mismatch [+definite, –unique], [–definite, +unique] ($6.46 - 0.0115 + 0.0285$). A magnified look at the post-spillover region gives us a clearer view that the advanced L2 speakers behave very similarly to L1 speakers.

As in L1 speakers, the interaction effect in the L2 data should be complemented with pairwise comparisons, which shows the imbalance between the two mismatching conditions. Figure 3.8 illustrates that there is an imbalance in terms of reading time increase in the two mismatching conditions. Figure 3.9 shows that the extent to which the pink and blue lines deviate from the black dot in the kitchen context is smaller than that in the appliance store context. Although the difference seems subtler compared to L1 data, the statistical analysis shows that L1 and L2 speakers pattern the same.

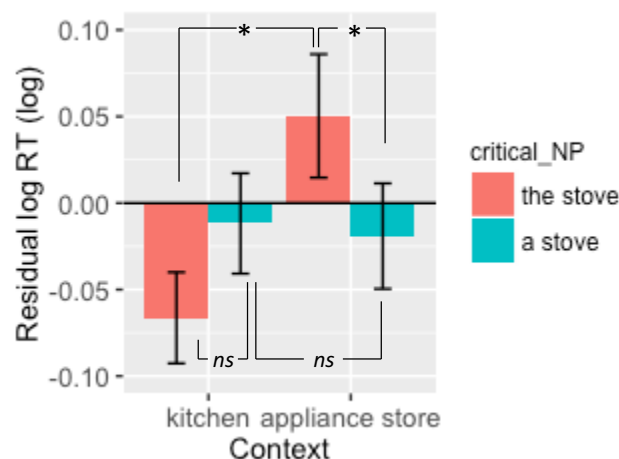


Figure 3.8. The interaction of uniqueness and definiteness in the post-spillover region (L2 data)
Note. The error bars indicate standard error.

The difference between the definite and indefinite conditions in the kitchen context is statistically insignificant ($\beta=-0.0068$, $SE=0.0288$, $t\text{-value}=-0.2354$, $p=0.8139$) while the difference in the appliance store is statistically significant ($\beta=-0.0639$, $SE=0.0282$, $t\text{-value}=2.2661$, $p=0.0234$). What should be noted here is how the blue line (indefinite) rarely moves when the context changes from the kitchen to the appliance store. It is in sharp contrast to the steep slope of the pink line (definite).

With the grouping factors changed (Figure 3.10), the discrepancy becomes more clearly visible and it is evident that indefinite NPs do not display a clear difference between the unique and non-unique contexts whereas definite NPs are matched with unique referents and mismatched with non-unique referents.

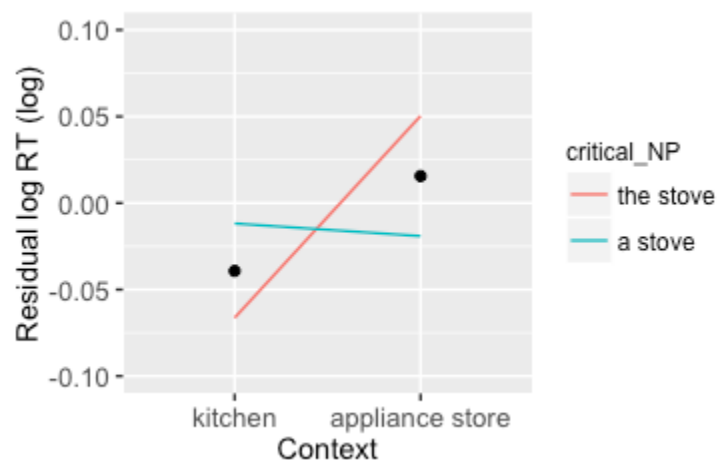


Figure 3.9. Comparison between definite and indefinite NPs by uniqueness.

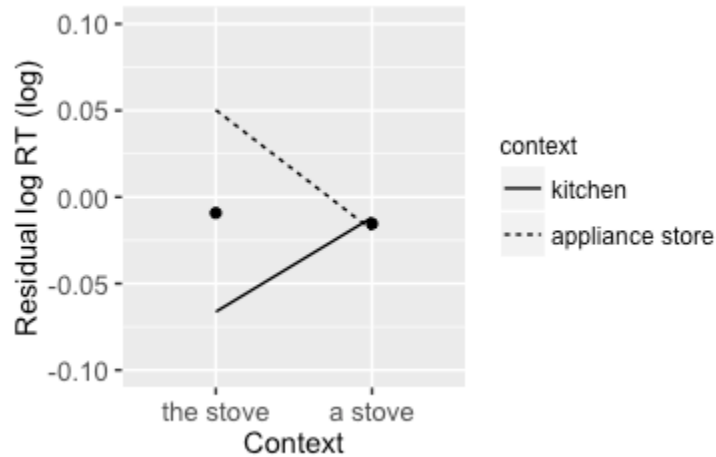


Figure 3.10. Comparison between the unique and non-unique referent contexts by definiteness

The results of L2 data analysis answer the same two research questions that the L1 results did. The interaction effect of uniqueness and definiteness is significant; definite and indefinite NPs are not significantly different in terms of processing cost in L2 processing as long as the presupposition of the (non-)uniqueness is in order. Pairwise comparisons show that there is an imbalance in the mapping between (in)definiteness and (non-)uniqueness.

The interpretation of the discrepancy has far greater implications for L2 English article use than in L1 processing. The following section discusses what the interaction effect and pairwise comparison results mean in L1 and L2 data and how they are related to the issues raised from the review of the existing literature.

3.6. Discussion

This chapter investigated L1 and L2 online sentence processing with regard to definiteness as uniqueness. The interaction effect of the presupposition of a (non-)unique referent

and the (in)definiteness of a referring NP is shown to be statistically significant, which was also reported in Clifton (2013). In addition, pairwise comparisons showed that processing cost increase incurred by the two mismatch conditions differ in terms of statistical significance both in L1 and L2 data. In L2, all such effects were observed in the post-spillover region, which was one region later than in L1. A plausible account for the delay is slower RTs in L2 speakers across all regions and conditions.

A direct comparison between the L1 speakers' RTs in the spillover region and the advanced L2 speakers' RTs in the post-spillover region might be questioned for its validity because the word lengths of the two regions are different and one is farther away from the critical region than the other. However, from the beginning of the sentence to the end, the L2 speakers were 70–150ms slower than the L1 speakers. Thus, the delayed effect could be due to the generally slower processing of L2 speakers. The striking similarity between the L1 spillover region and L2 post-spillover region can still be interpreted as evidence that such interactions exist in L2 processing as well. Thus, being slow should not mean that L2 speakers lack knowledge of (in)definiteness in English.

The interaction of uniqueness and definiteness has an important implication for both L1 and L2 processing of definiteness. The current study confirmed that definite and indefinite NPs newly introduced to the discourse do not differ in terms of processing cost unlike definite and indefinite NPs distinguished as given and new (Murphy, 1984). It also means definite and indefinite NPs are distinguished in terms of when it is more appropriate to use one or the other both in L1 and in L2 grammar.

A more important contribution of the current study is the pairwise comparisons between match and mismatch conditions made separately by context. The results of the comparisons can

be interpreted in two different ways. One is to translate the relatively less sensitive reaction to mismatched indefinite NPs in unique referent contexts in the current study to the overuse of *a(n)* in the previous research by Ionin et al. (2003, 2004). If such translation is possible, specificity can be seen as playing an important role in L2 English article use. Unfortunately, this argument is not decisive because the same relative lack of sensitivity was observed in L1 data as well. Adult natives' low sensitivity to mismatched indefinite NPs in unique referent contexts cannot be caused by their confusion between definiteness and specificity.

The other, more viable, interpretation is that the status of indefinite articles is not as clear-cut as definite articles to both L1 and L2 speakers. From the perspective of the mental discourse model, discussing a [+definite, +unique, –specific] referent in a conversation is a mutual agreement between interlocutors about a uniquely identifiable referent. In the example “In the kitchen, Jason checked out the stove very carefully,” the stove is very likely to have no specific referent (unless the communication occurred between two people who mutually know of the stove at issue¹³) but is agreed upon to exist in any given kitchen. When this presupposition is violated by an indefinite article, however, readers seem to accommodate the possibility that there might be more than one referents and that they are no longer unique.

Such pattern in both L1 and L2 could be due to the fact that an indefinite article has several functions. It could denote a certain referent (I have a car), it could denote a generic kind (I need a car), and it may be used to highlight singularity (I lost a shoe). And at the same time, it also indicates that the speaker does not presuppose the said entity to be uniquely identifiable to the hearer or mutually agreed upon between the interlocutors.

¹³ Given the nature of our lab experiment, such interpretation doesn't apply to the findings of the current study.

When this lack of significant reading time change is compared to the other mismatch case, it becomes clearer that the difference between the two mismatch conditions could well be explained by the nature of definite and indefinite articles. The most salient function of a definite article is its denotation of an entity that is (or is assumed to be) known to all conversation participants. This relatively powerful and clear function leads to the significant processing cost difference when they were used in the match and mismatch conditions.

In sum, the unbalanced processing cost increase in the two mismatch conditions points to more difficulty with indefinite articles than with definite articles in both language groups. The role of specificity account predicted the exact opposite pattern. Since the indefinite condition in the current experiment is [–definite, –specific] and the definite condition [+definite, –specific], the latter should have led to fluctuation and acquisition difficulty in L2 speakers. The prediction was not manifested in the data, which returns our discussion to the suspicion that the items used in the earlier studies could have induced an effect that may not exist.

This chapter investigated if L1 and L2 speakers were the same or different in processing definiteness as uniqueness. The answer should be both yes and no. Yes, they do pattern the same in mapping definite NPs to uniquely identifiable referents and in making a less clear distinction of (non-)uniqueness for indefinite NPs. But no, L2 speakers are not exactly the same as L1 speakers in that they reacted to experimental stimuli in general and to condition manipulations in a delayed manner. Some experimental design makes it harder for target-like L2 behavior to be detected (see Dekydtspotter, Schwartz, & Sprouse, 2006 for discussion). The delayed effects observed in the present study should not be disregarded as defects in L2 performance in that the quantitative difference should not be read as qualitative difference.

With the confirmation that advanced L2 speakers behave very much like L1 speakers in

processing definiteness as uniqueness, the next chapter will further explore the use of the definite and indefinite distinction to see how much predicting power the distinction has. Will both L1 and L2 speakers be able to use either a definite or an indefinite article to predict upcoming linguistic material? Chapter 4 uses a Referent Prediction Task to investigate this issue.

Chapter 4 Definiteness as a predictive cue for uniqueness

4.1 Introduction

The experiment described in Chapter 3 showed that advanced second language (L2) learners are sensitive to the mismatch between the uniqueness of a referent and the definiteness of a referring noun phrase (NP). However, that alone does not guarantee that L2 speakers will use definiteness information to predict what lies ahead in the sentence. The purpose of this chapter is to confirm that “definiteness” information within English articles provides a linguistic cue to predict a referent yet to be mentioned for both L1 and L2 speakers. A novel, ‘Referent Prediction Task’ was employed to probe the question. In what follows, the small number of studies that have considered the predictive role of English articles for unique and non-unique referents will be reviewed before the findings of the current research are described. This survey will be followed by the detailed description of the current experiment, whose results indicate that advanced L2 speakers have the same grammatical knowledge on English articles as L1 speakers. In addition, patterns observed in intermediate L2 speakers hint on the developmental path of English articles in L1 Korean learners of English.

4.2 Background

Various studies have looked into the semantics of definiteness using formal approaches (Abbott, 2006; Hawkins, 1991; Heim, 1982; Lyons, 1999; Russell, 1905 among many others). Researchers investigated the development of English articles from various perspectives in first

language (L1) acquisition (Bohnacker, 1997; Demuth, 2001a, 2001b; Gerken, 1991; 1994; 1996a, 1996b; Goad & Buckley, 2006; Guasti, Gavarró, de Lange, & Caprin, 2008; Tremblay, & Demuth, 2007; Wexler, 2011 among many others) and second language (L2) acquisition (Butler, 2002; Goad & White, 2009; Hopp, 2011; Huebner, 1979; 1983; Ionin, Baek, Kim, Ko, & Wexler, 2012; Ionin, Ko, & Wexler, 2004, 2007; Ionin, Zubizarreta, & Maldonado, 2008; Jarvis, 2002; Kim & Lakshmanan, 2009; Liu & Gleason, 2002; Parrish, 1987; Robertson, 2000; Sheen, 2007; Snape, Leung, & Ting, 2006; Trenkic, 2004, 2007, 2008; Trenkic & Pongpaiboj, 2013; Zdorenko & Paradis, 2008, 2011 among others). However, only a very few used psycholinguistic methods that focus on the prediction of what comes next within a sentence.

The studies previously mentioned in Chapter 3 (Burckhardt, 2006, 2008a, 2008b; Clifton, 2013; Murphy, 1984; Schumacher, 2009) used psycholinguistic experimental tasks. Although participants' sensitivity to pragmatic/semantic infelicitousness or grammatical anomalies is measured through the use of stimuli that violate their expectations (hence, predictions), it does not explicitly show which linguistic feature leads to the prediction of a particular referent in the visual scene. That is, participants' sensitivity to the mapping between (in)definite articles and (non-)unique referents can be observed through self-paced reading tasks and Event Related Potential experiments but they do not provide direct evidence about which article leads to the prediction of which referent.

Chambers, Tanenhaus, Eberhard, Filip, & Carlson (2002) and Trenkic, Mirkovic, and Altmann (2013), which was reviewed in Chapter 2, represent the closest attempts to observe L1 and L2 speakers' predictions on (non-)unique referents at the cue of (in)definite NPs. However, the two studies aimed to explore the effects of world knowledge in sentence processing and were not designed to tease apart grammatical knowledge and world knowledge. Because the aim of

the current chapter is to see the effect of grammatical knowledge only, an experiment designed solely for the effect of definiteness is needed.

Ahern and Stevens' (2014) experiment attempts to look into whether definite descriptions could lead to the prediction of unique referents in L1 sentence processing via a visual world paradigm eye-tracking experiment. Their experiments employed the notion of 'maximal uniqueness.'

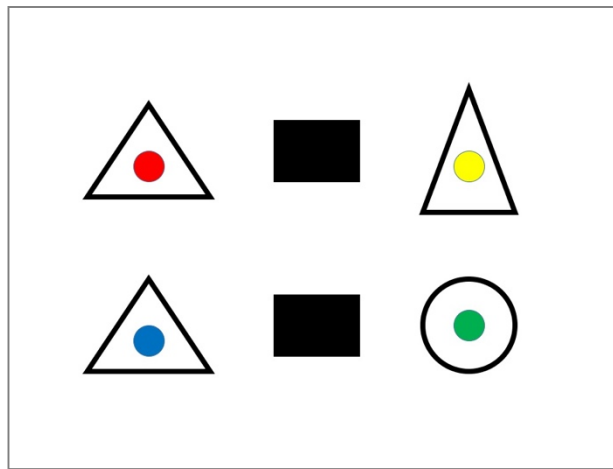


Figure 4.1. Visual stimuli re-produced from Ahern & Stevens (2014).

- (4.1)
- a. Click on the box that's next to the triangle with a yellow dot inside.
 - b. Click on the box that's next to the triangle with a red dot inside.
 - c. Click on the box that's next to a triangle with a yellow dot inside.
 - d. Click on the box that's next to a triangle with a red dot inside.

As is seen in Figure 4.1, all four figures are in a way unique. The three triangles are either different in shape or color. However, the two triangles on the left side of the figure are similar to each other and the one on the right-hand side with a yellow dot inside is distinct from the other

two. By constructing the visual stimulus this way, Ahern and Stevens (2014) could make the triangle with the yellow dot inside ‘a maximally unique referent.’ Of course, the maximally unique referent is a circle independent of the auditory stimuli in (4.1). However, once the critical NP is heard, the one with a yellow circle inside is maximally unique amongst the triangles.

The auditory stimuli in (4.1) are different from those in other studies (Chambers et al., 2002; Trenkic et al., 2013): the target NPs ‘a/the triangle with a yellow/red dot inside’ are embedded in a modifier clause instead of being used as the direct object of the preposition ‘on’ as in (4.2).

(4.2) ??Click on a triangle with a yellow dot inside.

Ahern and Stevens (2014) explain that it was to ensure that no auditory stimuli were pragmatically infelicitous. That design, however, prevented the authors from measuring eye gazes directly fixating on the target images on the screen. Because participants were instructed to click on the box, their eye gazes were not focused on the referents of the critical NPs. The results were obtained from combined looks to the entire row rather than to the individual objects and they used probability formula to calculate the advantage of the row with the maximally unique referent. However, the method and analysis were complicated and could only indirectly hint on the predictive nature of definite descriptions.

The goal of the current chapter is to determine whether both L1 and L2 speakers use definite and indefinite articles to predict unique and non-unique referents. Therefore, the experiment to be introduced tested only the predictive nature of articles without relying on the

interference of real-world knowledge, and predictions for individual objects were recorded rather than predictions for a set that includes predicted objects.

4.3 Method

4.3.1 Referent Prediction Task

A novel experimental method “referent prediction task” was devised to probe the question of whether definite and indefinite articles lead to unique and non-unique referents, respectively. The task makes use of auditory stimuli in the form of sentences with their last words omitted, as in (4.3), accompanied by a visual stimulus such as the one depicted in Figure 4.2. The hypothesis was that if the (in)definiteness of the article cues uniqueness at the end of the audio stimulus as in (4.3), participants will choose the unique referent in the given visual stimuli. The red cup and the three glasses are all unique on their own.

(4.3) The woman will buy the / a ...

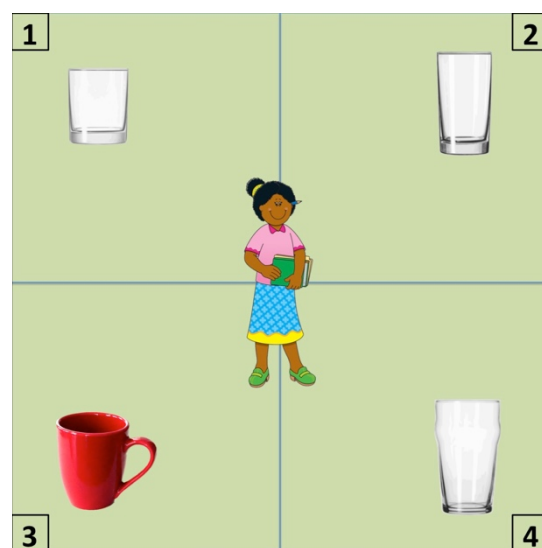


Figure 4.2. Visual stimuli for Referent Prediction Task

Although the three glasses are not identical to each other, the red cup does not share the property of transparency or the function that the three glasses share. This discrepancy makes the red mug maximally unique (Ahern & Stevens, 2014). Therefore, the combination of the audio and visual stimuli provides a good test case for the effect of definiteness in predicting unique (or non-unique) referents.

The benefit of the referent prediction task is that no infelicitous or grammatically anomalous stimuli are presented to participants. Ahern and Stevens (2014) also tried to avoid using pragmatically infelicitous or grammatically anomalous sentences by using NP modifiers “next to the...” One might ask why infelicitous or anomalous sentences should be avoided at all when it is common to compare grammatical and ungrammatical sentences (or felicitous and infelicitous sentences) in linguistic research. Below I will justify why avoiding infelicitous or ungrammatical constructions is necessary in the present research.

Experimental artifacts have been reported in the psycholinguistics literature that investigated referential expressions. In experiments where eye gaze is monitored to track the time-course of participants’ predictions for upcoming linguistic material, participants are exposed to semantically (or pragmatically) infelicitous and/or grammatically anomalous stimuli and get used to them. Then, they accommodate such sentences in later trials. This means that, towards the end of the experiment, expected patterns of behavior are harder to observe. Not surprisingly, Ito, Jincho, Minai, Yamane, and Mazuka (2012) reported a main effect of trial blocks.

If the current research question were investigated using eye-tracking, the auditory stimuli in (4.3) would be given as complete sentences as in (4.4). A visual-world eye-tracking paradigm compares the time it takes for participants’ eye gazes to locate a target in each condition.

However, constant exposure to conditions as in (4.4.b) and (4.4.c) would make participants expect infelicitous sentences, and so the effect of surprise would decrease in later trials.

- (4.4)
- a. The woman will buy the mug.
 - b. The woman will buy a mug.
 - c. The woman will buy the glass.
 - d. The woman will buy a glass.

That is, calling the red mug “a red mug” or one of the three glasses “the glass” with no further specification is semantically and pragmatically infelicitous in the visual context of Figure 4.2. But if participants are exposed to such sentences constantly, they will come to expect such sentences and their sensitivity to infelicitous sentences will not vary greatly by condition. By having participants choose one of the four referents as they hear a definite or an indefinite article, a referent prediction task more directly tests whether the presence of either article leads to the prediction of a different referent and avoids the infelicitousness issue.

Another advantage of using a referent prediction task is that the interpretation is straightforward. The interpretation of online measures can be nuanced. In recent decades, L2 researchers have strived to exploit psycholinguistic methods that probe Interlanguage without leading research participants to resort to their metalinguistic knowledge. However, online experimental methods do not solve all the problems. As shown in Chapter 3, L2 speakers can be slow in detecting grammatical anomalies in a self-paced reading task. Oftentimes, the slowness of L2 processing is misinterpreted as lack of knowledge because the effect does not appear in the same time frame as it would for L1 speakers. Or the effect spreads out to more than one region

and becomes harder to detect because processing is slower for L2 learners than for their native-speaker counterparts. However, slowness alone does not indicate that their processing mechanism is qualitatively different or that they lack the proper knowledge to detect grammatical anomalies. In this chapter, the referent prediction task is carried out under time pressure to prevent the exploitation of metalinguistic knowledge; at the same time, the task allows more straightforward and explicit interpretation.

4.3.2 Proficiency and participants

For proficiency, the same C-test (Appendix 1a) as in Chapter 3 (Experiment 1) was used. A total of 103 participants were recruited from the communities of Seoul National University in Seoul, Korea and the University of Hawai‘i at Mānoa in Honolulu, Hawai‘i. Table 4.1 summarizes the demographics of the participants, providing their age, gender, and C-test scores. The L2 speakers were limited to those born and raised in Korea.

Table 4.1. *Participant demographics*

		Age (years)	C-test (out of 40)	LOR (years)
L1 (n=33)	<i>mean</i>	26.62	33.98	
	<i>SD</i>	9.26	2.3	N/A
	<i>Range</i>	18–61	28–39	
Advanced L2 (n=42)	<i>mean</i>	23.74	32.21	0.55 year
	<i>SD</i>	3.39	2.5	1.28
	<i>Range</i>	19–32	29–38	0–6 years
Intermediate L2 (n=27)	<i>mean</i>	23.73	25	0.09 year
	<i>SD</i>	2.15	2.2	0.21
	<i>Range</i>	19–29	21–28	0–0.75

4.3.3 Stimuli

Audio and visual stimuli were used. A visual scene with a person or animal in the center and an object in each of the four corners was presented on a computer screen as shown in Figure 4.2. The four objects in the corners included one unique and three similar but not identical (thus, non-unique) items. The visual stimuli were presented 2 seconds prior to the auditory stimuli. This was to ensure that the participants were familiar with each object in the visual scene. Auditory stimuli were incomplete sentences that ended with either a definite or an indefinite article as in example (4.3).

The position of the unique referent was rotated around the four corners of the screen from the upper left-hand to the upper right-hand, to the lower left-hand, and to the lower right-hand corners (see Appendix 2a). A total of 20 critical items were presented in a Latin-square design in two separate lists where definite and indefinite conditions were alternated for each item. Another 40 items of a similar type were shuffled in as fillers to obfuscate the research purposes. Auditory stimuli in filler items did not include articles at the end of the sentence to ensure that participants' predictions in critical trials would not be affected by exposure to the combinations of articles and referents in filler trials. Also, the number of plausible referents in filler trials varied from one to all four objects on the screen to blur the distinction between critical trials (where there is only one unique referent in each trial) and filler trials and to prevent participants from selecting or avoiding unique referents in all trials.

4.3.4 Procedure

Participants were asked to fill out an online survey questionnaire on their language learning background before they came to the lab, which served to filter out people with language profiles unsuitable for the research purposes. Upon arriving at the lab, they were seated in front of a computer and listened to the instructions through a headset. Instructions, practice items, and experimental trials (including both critical and filler items) were presented via E-prime. The participants were given time to ask questions if they did not understand the task.

When a trial began, a 2-second preview of the visual scene (Figure 4.2) was given, and then the auditory stimuli (4.3) were presented. Participants were instructed to choose one of the four items to suit the sentence as the last omitted word. Each incomplete sentence was immediately followed by a beep; two consecutive beeps followed five seconds after the first beep. Participants were instructed to press the button as fast as possible but only after they heard the first beep and before they heard the last two beeps.

This design had two important purposes. One was to ensure that they were using article information to predict upcoming material. If they pressed the button before they heard the article just to meet the requirement of “pressing the button as fast as possible,” the results would not have been able to show the effects of (in)definite articles. The other purpose was to prevent participants from resorting to their metalinguistic knowledge. Giving them enough time to ponder their choices could lead to their trying to figure out the research purposes or the linguistic patterns of the items. At the analysis stage, responses were checked to make sure no button presses occurred earlier than exposure to the articles. No data points were eliminated from overly fast button presses and less than 1% of the data were discarded due to overly slow button presses (RTs longer than 3 standard deviations). The experiment took around 15 minutes for each participant, and the C-test that followed generally took another 15 minutes.

4.4 Analysis

The prediction of unique vs. non-unique referents. When a unique referent was predicted, the response was coded as 1; when one of the non-unique referents was predicted, it was coded as 0. Because the dependent measure was binary, a logit mixed effects model (i.e., a generalized linear mixed effects model for binomially distributed outcomes) was used (Jaeger, 2008).

Analyses were run in R (R Core Team, 2016) using the lmer package (Bates, Maechler, Bolker, & Walker, 2015). A random intercepts model was fit with subjects and items as random effects for all analyses. First, each group's choice of either unique or non-unique referents was separately modeled with article as a fixed factor (4.3). Later, the three groups were all modeled together to see if there was any significant interaction effect of article and speaker group (4.4). For the interaction of article by group, not all models merged with a maximal random effect structure. According to Bates, Kliegl, Vasishth, and Baayen (2015)¹⁴, an lmer model with a maximal random effect structure might not merge when the model lacks enough observations given the complexity of the model. In such cases, simpler models were selected using Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC)¹⁵ following Bates et al. (2015). The model of L1 and advanced L2 comparison merged with the random slopes of subject and item for article (4.4a–4.4b). The models that compared L1 vs. intermediate L2 and advanced vs. intermediate L2 did not converge with random slopes; hence, the latter include only random intercepts (4.4c–4.4d). Finally, the model that compared the intermediate L2 group to the rest of the participants (L1 and advanced L2) merged with only the random intercept of subject and the

¹⁴ Cf. Barr, Levy, Scheepers, and Tily (2013).

¹⁵ AIC indicates the information loss over a given data set and BIC mitigates the risk of overfitting. For both, lower values indicate better models.

random intercept and slope of item (4.4e–4.4f). (4.3) and (4.4) show glmer (generalized linear mixed effects regression) R codes and their corresponding mathematical equations used for data analysis.

```
(4.3a) glmer(def~article+(1|Subject)+(1|item), data, family = binomial)
```

$$(4.3b) Y_{si} = \beta_{00} + \beta_1 * X + \gamma_{0s} + \delta_{0i} + e_{si}$$

Y_{si} is the response for subject s in item i ;

X is the fixed factor taking values of 1 (=definite article) and -1 (=indefinite article) depending on its experimental condition;

β_{00} is the grand mean of the responses;

β_1 is the regression coefficient relating X to the Y_{si} , in this case, for the main effect of the fixed factor ‘article’;

γ_{0s} is the subject intercept, the deviation from β_{00} for subject s ;

δ_{0i} is the item intercept, the deviation from β_0 for item i ;

e_{si} is the observation-level error with its mean 0 and variance σ^2 .

```
(4.4a) glmer(def~article*group+(1+article|Subject)+(1+article|item), data,
family = binomial)
```

$$(4.4b) Y_{si} = \beta_{00} + \gamma_{0s} + \delta_{0i} + (\beta_1 + \gamma_{1s} + \delta_{1s}) * X + \beta_2 * W + \beta_3 * XW + e_{si}$$

Y_{si} is the response for subject s in item i ;

X_i is the fixed factor ‘article’ taking values of 1 (=definite) and -1 (=indefinite);

W_i is the fixed factor ‘group’ taking values of 1 (=native) and -1 (=adv L2);

- β_{00} is the grand mean of the responses;
- γ_{0s} is the subject intercept, the deviation from β_{00} for subject s ;
- δ_{0i} is the item intercept, the deviation from β_0 for item i ;
- β_1 is the regression coefficient relating X to the Y_{si} , in this case, for the main effect of the fixed factor ‘article’;
- γ_{1s} is the subject slope with regard to the fixed factor X ;
- δ_{1i} is the item slope with regard to the fixed factor X ;
- β_2 is the regression coefficient relating W to the Y_{si} , in this case, for the main effect of the fixed factor ‘group’;
- β_3 is the regression coefficient relating XW to the Y_{si} ;
- e_{si} is the observation-level error with its mean 0 and variance σ^2 .

```
(4.4c) glmer(def~article*group+(1|Subject)+(1|item), data, family = binomial)
```

$$(4.4d) Y_{si} = \beta_{00} + \gamma_{0s} + \delta_{0i} + \beta_1 * X + \beta_2 * W + \beta_3 * XW + e_{si}$$

- Y_{si} is the response for subject s in item i ;
- X_i is the fixed factor ‘article’ taking values of 1 (=definite article) and -1 (=indefinite article);
- W_i is the fixed factor ‘group’ taking values of 1 and -1 (for native and low L2; for adv L2 and low L2);
- β_{00} is the grand mean of the responses;
- γ_{0s} is the subject intercept, the deviation from β_{00} for subject s ;
- δ_{0i} is the item intercept, the deviation from β_0 for item i ;

- β_1 is the regression coefficient relating X to the Y_{si} , in this case, for the main effect of the fixed factor ‘article’;
- β_2 is the regression coefficient relating W to the Y_{si} , in this case, for the main effect of the fixed factor ‘group’;
- β_3 is the regression coefficient relating XW to the Y_{si} .
- e_{si} is the observation-level error with its mean 0 and variance σ^2 .

```
(4.4e) glmer(def~article*group+(1|Subject)+(1+article*group|item), data,
            family = binomial)
```

$$(4.4f) \quad Y_{si} = \beta_{00} + \gamma_{0s} + \delta_{0i} + (\beta_1 + \delta_{1i}) * X + (\beta_2 + \delta_{2i}) * W + (\beta_3 + \delta_{3i}) * XW + e_{si}$$

- Y_{si} is the response for subject s in item i ;
- X_i is the fixed factor ‘article’ taking values of 1 (=definite article) and -1 (=indefinite article);
- W_i is the fixed factor ‘group’ taking values of 1 (= low L2) and -1 (=rest of the participants);
- β_{00} is the grand mean of the responses;
- γ_{0s} is the subject intercept, the deviation from β_{00} for subject s ;
- δ_{0i} is the item intercept, the deviation from β_0 for item i ;
- β_1 is the regression coefficient relating X to the Y_{si} , in this case, for the main effect of the fixed factor ‘article’;
- δ_{1i} is the item slope with regard to the fixed factor X ;

- β_2 is the regression coefficient relating W to the Y_{si} , in this case, for the main effect of the fixed factor ‘group’;
- δ_{2i} is the item slope with regard to the fixed factor W ;
- β_3 is the regression coefficient relating XW to the Y_{si} .
- δ_{3i} is the item slope with regard to the interaction of XW ;
- e_{si} is the observation-level error with its mean 0 and variance σ^2 .

4.5 Results

The analysis of the data resulted in the same pattern of behavior in L1 and advanced L2 speakers and the opposite pattern of behavior in intermediate L2 speakers. As Figure 4.3 shows, all groups showed a bias towards a unique referent. Regardless of definiteness, they chose a unique referent more than 50% of the time. However, what’s important is that the (in)definiteness cue provided at the end of the incomplete sentences made a significant difference in the selection of either unique or non-unique referents both for L1 speakers and for advanced L2 speakers. These two groups chose a unique referent significantly more often when a definite article was provided (solid line) and the main effect of article was significant in both groups (L1: $p = .022$; advanced L2: $p = .008$). That is, when the conditions changed from definite to indefinite article, the rate of their unique referent selection went down significantly. The pattern that the intermediate L2 speakers showed is worth noting. They chose a unique referent numerically more often when an indefinite article was given as a cue, than when a definite article was given. In sections 4.5.3 and 4.5.4, I will elaborate on the possible reasons why the difference

was only numerical and not statistically significant and explain why even the numerical difference should be meaningfully interpreted with reference to the other groups of speakers.

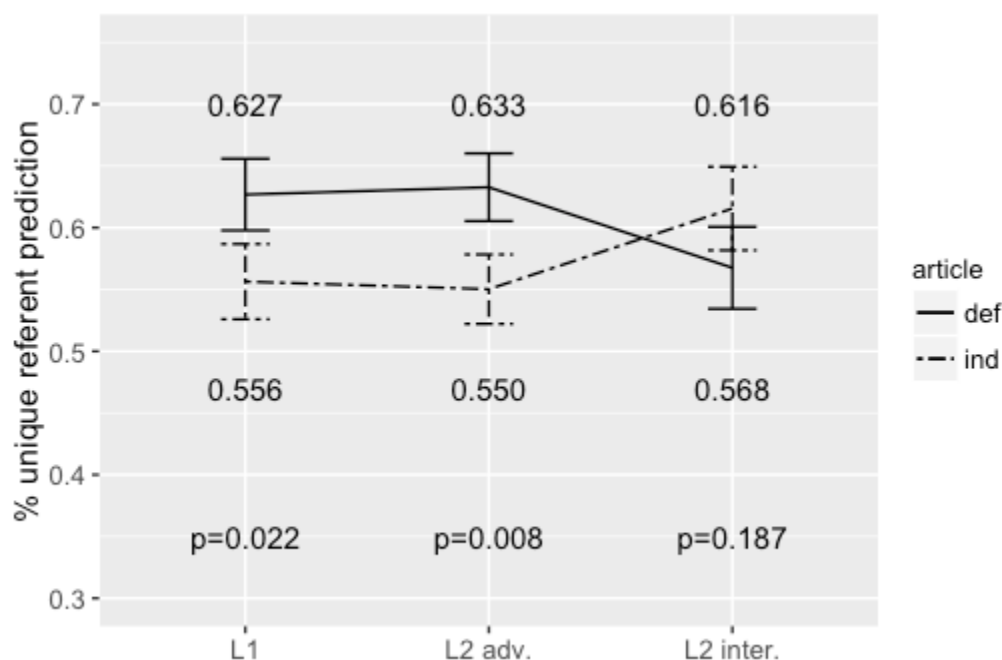


Figure 4.3. Ratio of unique referent prediction by speaker group and by article.

Note. Error bars indicate standard errors.

4.5.1 L1 speakers

Table 4.2. summarizes the fixed effect of article in L1 data. The glmer model was fit with sum coding and the output table can be understood in the same way the output table was interpreted in Chapter 3. The intercept is the grand mean and the other row indicates the main effect of article. The grand mean on the intercept is the average of all data points in L1. That is, regardless of which article is given, L1 participants will predict unique referents 61.8% of the time. This means a strong bias towards a unique referent. The second row ‘article’ indicates the

amount of change in estimates created by the choice of article (*a* or *the*). So, the absolute value of change in estimates is 20.9%.

Table 4.2. *The glmer output for the main effect of article in L1 speakers.*

(a) Fixed effect

		Estimate	Std. Error	z value	p value	
(Intercept)	β_{00}	0.6179	0.2635	2.3450	0.0190	*
article	β_1	0.2085	0.0913	2.2850	0.0223	*

· $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$

(b) Random effects

Groups	Name	Variance	Std.Dev.
Subject	(Intercept)	γ_{00}	0.6233
item	(Intercept)	δ_{00}	0.8333

Number of obs: 654, groups: Subject, 33; item, 20

As is seen in Figure 4.3, the definite condition elicits higher prediction for unique referents. Therefore, it will increase the estimate to 82.7% (=61.8+20.9) and the indefinite condition will decrease the estimate to 40.9% (=61.8–20.9). A caveat in comparing the glmer output of the fixed effect and the visualization of the actual data is that the glmer fixed effect output is a predictive model based only on the fixed factor (article) and the line graph of the actual data includes all random variances resulting from individual differences among subjects and items. That is, the coefficients in the output tables indicate the effect of definite and indefinite articles on predicting unique referents assuming no other influence from external factors.

In brief, L1 speakers do use the linguistic information of definite and indefinite articles to predict an unmentioned referent. Their prediction of a unique and non-unique referent is

dependent on the article they hear at the end of the auditory stimuli and the difference between the two conditions is statistically different.

4.5.2 Advanced L2 speakers

Table 4.3 shows the output for advanced L2 speakers and it can be understood the same way Table 4.2 was interpreted. Compared to L1 speakers, whose grand mean was 61.8%, advanced L2 participants' bias towards the unique referent was not as high; their grand mean is 50.3%.¹⁶ However, the main effect of article is 20.3%, which is approximately the same as L1 counterparts.

Table 4.3. *The glmer output for the main effect of article in advanced L2 speakers.*

(a) Fixed effect

		Estimate	Std. Error	z value	p value	
(Intercept)	β_{00}	0.5032	0.1911	2.6330	0.0085	**
article	β_1	0.2027	0.0766	2.6450	0.0082	**

· $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$

(b) Random effects

Groups	Name		Variance	Std.Dev.
Subject	(Intercept)	γ_{00}	0.3900	0.6245
item	(Intercept)	δ_{00}	0.4234	0.6507

Number of obs: 840, groups: Subject, 42; item, 20

¹⁶ The grand mean of 50.3% might be perceived as no bias at all; however, all trials included four referents (a single unique one and three similar ones). Mathematically, each referent should have a 25% chance of being selected. Even after allowing for the similarity of the three non-unique referents, a 50% probability of predicting a unique referent should not be interpreted as no bias.

This means that advanced L2 speakers will predict a unique referent 50% of the time but a definite article increases the probability 20.9% while an indefinite article will decrease the probability by the same amount. In other words, the probability that advanced L2 speakers will choose a unique referent goes up to 70.6% in the definite condition but down to 30% at the cue of an indefinite article. Again, the predictive model in the fixed effects table does not correspond 100% to the observed values since random variances are presented separately in the random effects table. However, the z and p values indicate their statistical significance and the observed effects are not merely due to random variances in individual subjects' random behavior or experimental items' peculiarity.

4.5.3 Intermediate L2 speakers

Table 4.4 reveals that intermediate L2 speakers do not show a great bias towards unique referents at all. That is, the grand mean (intercept) was only .230 and its p value was not significant. This means the coefficient estimate .230 for the intercept was not meaningfully different from 0. Also, the coefficient estimate of $-.130$ indicates the reverse pattern that intermediate L2 speakers show in reaction to definite articles. That is, unlike L1 and advanced L2 speakers who showed an increased rate of unique referent prediction at the cue of a definite article, the intermediate L2 group showed a decreased rate of unique referent prediction.

Based on the coefficient estimate for the fixed factor 'article' in Table 4.4., the likelihood that intermediate L2 speakers will predict a unique referent at the cue of a definite article is 10% ($=23-13$). Likewise, the likelihood they will predict a unique referent at the cue of an indefinite article is 36% ($=23+13$). Again, the coefficient values on the fixed effects output table shows the

effect of fixed factors only; article in this case. Therefore, the predicted coefficient cannot be the same as the observed value that includes random variances from both subjects and items.

Table 4.4. *The glmer output for the main effect of article in intermediate L2 speakers.*

(a) Fixed effect

		Estimate	Std. Error	z value	p value
(Intercept)	β_{00}	0.2295	0.2669	0.8600	0.3900
article	β_1	-0.1303	0.0987	-1.3200	0.1870

$p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$

(b) Random effects

Groups	Name	Variance	Std.Dev.
Subject	(Intercept)	γ_{00} 0.3438	0.5864
item	(Intercept)	δ_{00} 0.9739	0.9869

Number of obs: 539, groups: Subject, 27; item, 20

What is noteworthy here is the reversed pattern observed in the intermediate L2 group. As the overlapping error bars in Figure 4.3 and insignificant p values in Table 4.4 indicate, the difference between the two conditions within the intermediate group itself may not seem very important. The lack of statistical significance, however, should not be a reason to disregard the opposite pattern shown by the intermediate group. This insignificance could be due to a smaller number of observations compared to L1 and advanced L2 speakers where data were collected from 33 and 42 participants, respectively. The number of intermediate L2 participants was only 27. It remains as a topic for further study if a larger sample size will increase the statistical significance. The implication of intermediate learners' behavior will be further discussed in the next section when group comparisons shed light on the interaction effect of article and speaker group.

4.5.4 Interaction of article and speaker group in group comparisons

L1 vs. Advanced L2. What should be carefully considered is the interaction effect of article and speaker group. Table 4.5 is the output summary of a glmer model that fits the interaction between article and speaker group in L1 and advanced L2 speakers. The grand mean of unique referent prediction is about 56%, which roughly matches the average of the grand mean of the two groups: 61.8% in L1 speakers and 50.3% in advanced L2 speakers.

Table 4.5. *The glmer output for the interaction effect of article and group in L1 and adv L2.*

(a) Fixed effects

		Estimate	Std. Error	z value	p value	
(Intercept)	β_{00}	0.5599	0.1925	2.9090	0.0036	**
article (X)	β_1	0.2003	0.0601	3.3310	0.0008	***
group (W)	β_2	0.0506	0.0983	0.5150	0.6064	
article x group (XW)	β_3	-0.0095	0.0585	-0.1630	0.8708	

· $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$

(b) Random effects

Groups	Name	Variance	Std.Dev.	Corr
Subject	(Intercept)	γ_{00}	0.4681	0.6841
	article	γ_1	0.0010	0.0316
item	(Intercept)	δ_{00}	0.5438	0.7374
	article	δ_1	0.0003	0.0165

Number of obs: 1514, groups: Subject, 76; item, 20

The coefficient estimate for the article is .200, which is also the same as that in L1 and advanced L2 speakers. Because both L1 and advanced L2 speakers showed a significant increase of unique referent prediction in the definite condition, they show the same significant main effect

of article; thus, no main effect of speaker group was observed. The two groups behaved alike and no significant interaction effect of article and speaker group were observed.

L1 vs. intermediate L2. Table 4.6 shows the output of a glmer model that includes only L1 speakers and intermediate L2 speakers. This model shows a marginal main effect of group and a significant interaction effect of article and group. The main effect of article observed in L1 speakers and advanced L2 speakers cannot be seen here. This is due to the opposite patterns of behavior in the two groups. That is, the unique referent prediction rate increases in L1 but decreases in intermediate L2 at the cue of a definite article; thus, a main effect observed in the L1 group analysis is canceled out by the opposite behavior of the intermediate L2 speakers.

Table 4.6. *The glmer output for the interaction effect of article and group in L1 and inter L2.*

(a) Fixed effects

		Estimate	Std. Error	z value	p value
(Intercept)	β_{00}	0.4196	0.2193	1.9140	0.0557
article (X)	β_1	0.0272	0.0655	0.4160	0.6776
group (W)	β_2	0.2102	0.1087	1.9350	0.0530
article x group (XW)	β_3	0.1625	0.0657	2.4720	0.0134 *

· $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$

(b) Random effects

Groups	Name	Variance	Std.Dev.
Subject	(Intercept)	γ_{00} 0.4515	0.6719
item	(Intercept)	δ_{00} 0.7230	0.8503

Number of obs: 1213, groups: Subject, 61; item, 20

Figure 4.4 clearly visualizes why the main effect of article cannot be observed here. The dotted line hiding behind the solid line in Figure 4.4 is the L1 group. And the dashed line in the opposite direction is the intermediate group. Because the two groups make different

predictions for each article, the main effect of article cannot be observed. Instead, the marginal main effect of group indicates that the overall rate at which L1 speakers predict unique referents was higher than intermediate L2 learners would do so.

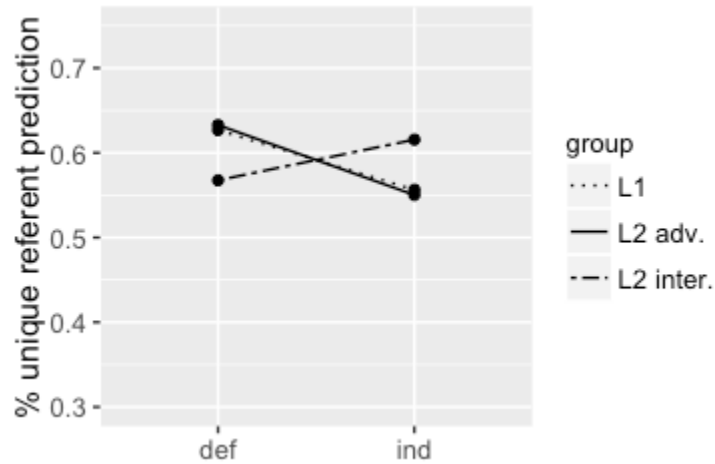


Figure 4.4. The interaction of speaker group and article.

Finally, the cross-over interaction effect is significant at the .013 alpha level. This means that the opposite tendency observed between the two groups is not a mere coincidence. It is for this reason that the numerical difference between definite and indefinite conditions observed in the intermediate L2 group analysis cannot be disregarded. Although the pairwise comparison between definite and indefinite conditions within the intermediate L2 group was not statistically significant, the interaction effect clearly shows how the linguistic information (definite vs. indefinite articles) takes a different effect by speaker group.

Advanced L2 vs. intermediate L2. The comparison between advanced L2 speakers and intermediate L2 speakers showed the same result. Table 4.7 summarizes the output for the glmer model with only L2 speakers. There was no main effect of article for the same reason as in the

comparison between L1 and intermediate L2 speakers. Because advanced L2 speakers predicted a unique referent more at the cue of a definite article and intermediate L2 speakers did so at the cue of an indefinite article, no main effect of article was observed across the two groups. No main effect of group was observed, which means that, unlike in the comparison between L1 and intermediate L2, the grand mean difference between advanced and intermediate L2 was not statistically significant. The cross-over interaction of article and speaker group was significant. This means that the opposite behavior of the two speaker groups was due to chance.

Table 4.7. *The glmer output for the interaction of article and group between adv. and inter. L2.*

(a) Fixed effects

	Estimate	Std. Error	z value	p value	
(Intercept)	β_{00} 0.3707	0.2018	1.8370	0.0662	.
article (X)	β_1 0.0276	0.0614	0.4500	0.6528	
group (W)	β_2 0.1537	0.0969	1.5860	0.1128	
article x group (XW)	β_3 0.1769	0.0617	2.8690	0.0041	**

. $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$

(b) Random effects

Groups	Name	Variance	Std.Dev.
Subject	(Intercept)	γ_{00} 0.3703	0.6085
item	(Intercept)	δ_{00} 0.6239	0.7899

Number of obs: 1379, groups: Subject, 69; item, 20

Intermediate L2 vs. the rest. Since L1 and advanced L2 speakers showed an identical behavior, another glmer model was fit where all three speaker groups were included. In this model, L1 and advanced L2 speakers were collapsed into one group. The interaction effect of article and speaker group was significant in this model as well. The output summary is provided in Table 4.8. This also indicates that the opposite pattern observed in intermediate L2 speakers

cannot be dismissed just because the main effect of article within the intermediate group was not significant.

Table 4.8. *The glmer output for the interaction of article and group b/w inter. L2 vs. the rest*

(a) Fixed effects

	Estimate	Std. Error	z value	p value
(Intercept)	β_{00} 0.3942	0.2114	1.8650	0.0622
article	β_1 0.0316	0.0602	0.5240	0.6003
group	β_2 -0.1608	0.1072	-1.5000	0.1337
article x group	β_3 -0.1651	0.0580	-2.8460	0.0044

· $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$

(b) Random effects

Groups	Name	Variance	Std.Dev.	Corr
Subject	(Intercept)	γ_{00} 0.4310	0.6565	
item	(Intercept)	δ_{00} 0.7178	0.8473	
	article	δ_1 0.0059	0.0770	0.10
	group	δ_2 0.0552	0.2350	0.57 0.88
	Article x group	δ_3 0.0005	0.0220	0.39 0.96 0.98

Number of obs: 2053, groups: Subject, 103; item, 20

Their pattern of behavior, which is in the opposite direction from the other groups, is noteworthy enough due to this interaction effect. This interaction indicates how the same linguistic information takes a completely different effect in the speaker groups and the different patterns of behavior might signify the developmental path of English article acquisition in L1 Korean L2 English speakers.

One possible interpretation of the intermediate L2 speakers' behavior is that they might interpret the indefinite article as a marker of singularity. Most English-Korean dictionaries list 'one' as the first entry for the meaning of 'a(n),' and sample sentences also have direct

translations including the word *hana* in Korean meaning ‘one.’ Due to the influence of their L1 or their English instruction in Korea, Korean learners of English might think of the indefinite article as a singularity marker. If this is the case, the effect of article in the intermediate group might have been significant if the study had included a larger sample size; this question remains for further research.

4.6 Discussion

The goal of the chapter was to determine whether L1 and L2 speakers of English share the same linguistic information with regard to English articles. The referent prediction task tested whether definite and indefinite articles lead to the prediction of unique and non-unique referents, respectively. The findings can be summarized as the following three: (1) There was an overall bias towards a unique referent regardless of given articles, (2) L1 and advanced L2 speakers predicted unique referents significantly more when the auditory stimuli ended with a definite article than with an indefinite article, and (3) intermediate L2 speakers showed the tendency of predicting a unique referent at the cue of an indefinite article more than with a definite article, which might indicate the use of *a/an* as a singularity marker.

The overall bias towards unique referents could be due to the experimental design. The referents are always grouped into one unique referent and three non-unique referents. Therefore, the unique referent always stands out, which might have drawn participants’ attention. Also, with proper modification, the unique referent could be definite but it could be indefinite as well. That is, the mug in Figure 4.1 could be referred to as ‘a mug’ in a sentence like “I see a mug in the bottom left corner.” A non-unique referent could become unique when restrictive modifiers are

added such as “I will predict the glass in the upper right-hand corner.” The mapping between (in)definite articles and (non-)unique referents are thus not strictly dichotomous. A strong bias toward the unique referent could, therefore, arise because of its visual salience regardless of whether the sentence ended with either with *the* or *a*.

Despite the overall bias, the statistical analysis shows a significant difference in terms of choosing a unique referent between the definite and indefinite conditions. This difference was observed in L1 and advanced L2 speakers only, which was the main goal of the current experiment. To compare how linguistic and non-linguistic information is integrated in L1 and L2 speakers, it needed to be determined that both groups grasp the essential of the definiteness contrast expressed by English articles—a type of grammatical knowledge. The results show that at least advanced L2 speakers behave the same as L1 speakers in that both groups predicted unique referents significantly more when definite articles were given than indefinite articles.

The results dovetail with what was observed in the previous chapter and also with the findings in Trenkic et al. (2013), reviewed in Chapter 2. The self-paced reading experiment in the previous chapter and the eye-tracking experiment in Trenkic et al. (2013) had different research questions and experimental methods; however, they both indicated that L2 speakers use the (in)definiteness distinction of English articles to refer to unique and non-unique referents in a native-like manner.

Finally, it is noteworthy that the pattern observed in the intermediate L2 speakers was the opposite from the other groups. This discrepancy seems to delineate the developmental trajectory of English articles in Korean learners. With the influence of their L1 in the initial state, they use indefinite articles to mark singularity but as their proficiency level improves, they develop target-

like use of English articles to distinguish definiteness (manifested as unique identifiability in the current study).

In conclusion, the results of Chapter 3 and the current chapter indicate that advanced L2 speakers have the same grammatical knowledge as L1 speakers when it comes to using (in)definite articles to predict (non-)unique referents. The following chapter will test whether the processing of linguistic information is affected by the presence of non-linguistic information.

Chapter 5 The integration of linguistic and non-linguistic information in L2 processing

5.1 Introduction

The purpose of this chapter is to observe the behavior of first language (L1) and second language (L2) speakers when two different sources of information are provided at the same time: linguistic and non-linguistic information. It was already confirmed in previous chapters that linguistic information is shared between L1 and advanced L2 speakers: the same grammatical knowledge of English articles as markers of ‘definiteness,’ operationalized as unique identifiability. In Chapter 3, a self-paced reading task showed that both L1 and advanced L2 speakers have the sensitivity to the mapping between unique referents and definite NPs although the latter were slower than the former. Chapter 4 also showed that both L1 and advanced L2 speakers use English articles as linguistic cues to predict unique vs. non-unique referents. Definite articles led to the prediction of unique referents significantly more than indefinite articles did. It is also important to remember that, in Chapter 4, intermediate L2 speakers showed the pattern of using indefinite articles as markers of singularity. The current chapter reports the results of a norming study on real-world knowledge and the results of an integration experiment in which both linguistic and non-linguistic information are presented within a sentence.

The findings suggest that L1 speakers integrate both linguistic and non-linguistic information incrementally, which means that they generate predictions for upcoming linguistic material and revise their predictions using both linguistic and non-linguistic information. Also, L1 speakers could respond to both linguistic and non-linguistic information at the same rate. On the other hand, advanced L2 speakers could not incorporate linguistic information in a native-

like manner. Although L2 speakers could incorporate both sources of information in sentence processing, processing non-linguistic information preceded processing linguistic information. The results are discussed in light of how processing efficiency can account for the L2 speakers' behavior.

5.2 Background

As was discussed in Chapter 2 and Chapter 4, attempts to investigate the role of real-world knowledge in sentence processing is not new (Chambers, Tanenhaus, Eberhard, Filip, & Carlson, 2002; Martin, Garcia, Breton, Thierry, & Costa, 2014; Martin, Garcia, Breton, Thierry, & Costa, 2015; Trenkic, Mirkovic, & Altmann, 2013). However, the focus of these studies have not been on teasing apart grammatical knowledge and real-world knowledge when they are co-present in a sentence.

Figure 5.1 is an example of visual stimulus in Chambers et al. (2002), which was accompanied by auditory stimuli as in (5.1). In the visual scene (Figure 5.1), either both cans (the goal objects) in the visual scene were plausible targets or only one of them could be a target depending on the size of the cube (the theme object). If the large cube was given as the theme object, only the large can could be the goal; then, it was the one-referent condition. When the theme (cube) was small, both cans could be goals; then, the visual scene became the two-referent condition. In the one-referent condition, the goal object is unique while there are two, non-unique, goals in the two-referent condition. The former aligns with a definite NP *the can* (5.1a) and the latter with an indefinite NP *a can* (5.1b).

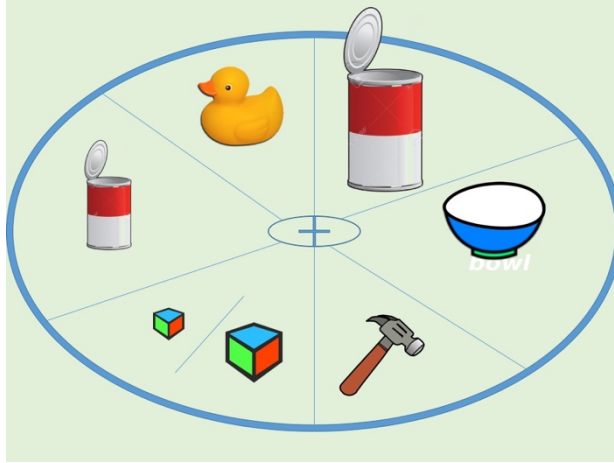


Figure 5.1. Visual scene sample re-created from Chambers et al. (2002); Second experiment

- (5.1) a. Put the cube inside the can.
b. Put the cube inside a can.

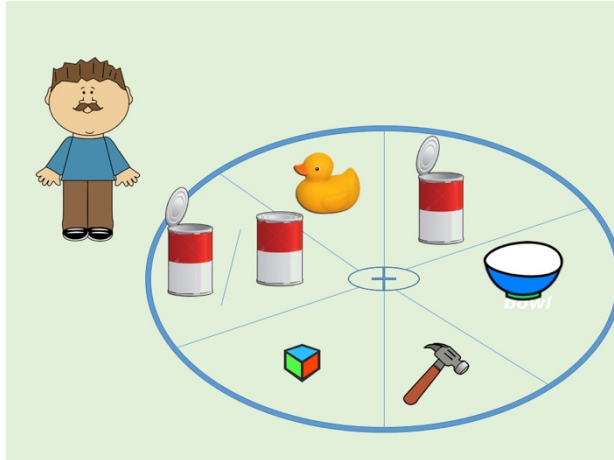


Figure 5.2. Visual stimulus re-produced from Trenkic et al. (2013).

- (5.2) a. The man will put the cube inside the can.
b. The man will put the cube inside a can.

Likewise, Trenkic et al. (2013) used almost the same idea except that the world knowledge involved whether both cans were open (Figure 5.2) rather than manipulating the size of containers and theme objects. When an open can and a closed can were given in the visual scene, it was the one-referent condition, which requires a definite NP (5.2a). When two open cans were presented, it was the two-referent condition, which aligns with an indefinite NP (5.2b).

The results of both experiments were interpreted that L1 (Chambers et al., 2002) and L2 (Trenkic et al., 2013) speakers could incorporate real-world knowledge incrementally in real-time. However, the experiments did not include conditions in which linguistic knowledge and non-linguistic knowledge pointed to different referents and created conflicts between the two sources of information. To see the effects of linguistic and non-linguistic information independently, the current chapter employs a Referent Identification Task that measures participants' reaction time to the accordance and discordance of auditory and visual stimuli. Figure 5.3 is an example where the visual stimulus includes real-world knowledge that the person in the picture (i.e., a doctor) is more strongly associated with the stethoscope than to the other objects).

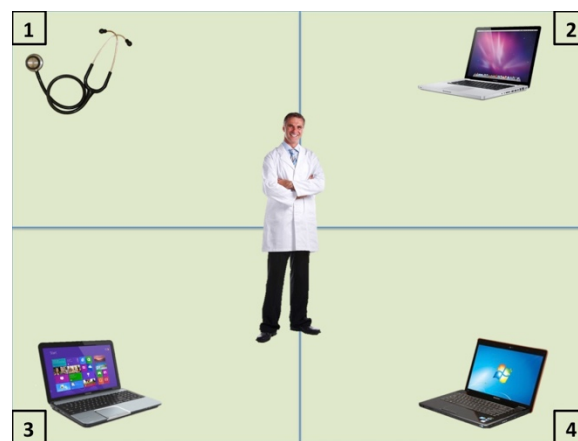


Figure 5.3. A visual stimulus example in which world knowledge predicts a unique referent

- (5.3) a. The man will want to use **the** stethoscope.
b. The man will want to use **a** stethoscope.
c. The man will want to use **the** laptop.
d. The man will want to use **a** laptop.

If such a visual scene is accompanied by auditory stimuli as in (5.3), will the linguistic information embedded in the definite and indefinite articles (bolded) in the critical noun phrase affect the way L1 and L2 speakers process the sentences? Real-world knowledge points to the stethoscope, a unique referent that requires a definite description. An indefinite NP *a stethoscope* satisfies one's expectation based on world knowledge but not based on grammatical knowledge.

Real-world knowledge might point to non-unique referents as in Figure 5.4. Because of her lab coat, the person in the center seems to be more obviously associated with microscopes than a sweater. Anyone can wear a sweater so it is not a completely implausible option.¹⁷ When world knowledge points to non-unique referents and participants predict one of the microscopes, an indefinite description (i.e., *a microscope*) suits the visual scene better than a definite description. The auditory stimuli used for Figure 5.4 are given in (5.4).

- (5.4.) a. She will want to get **a** microscope.
b. She will want to get **the** microscope.
c. She will want to get **a** sweater.
d. She will want to get **the** sweater.

¹⁷ Likewise, laptops are not completely implausible options for the doctor in Figure 5.3.

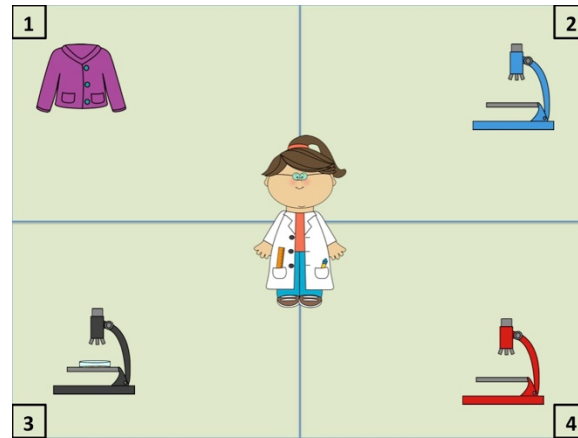


Figure 5.4. A visual stimulus example in which world knowledge predicts non-unique referents

In both (5.3) and (5.4), two sentences are grammatical while the other two are ungrammatical. Also, two sentences match real-world knowledge and the other two do not. When grammatical knowledge and real-world knowledge conflict with each other, how will this be processed by L1 and L2 speakers react to this in terms of their processing? The studies reviewed in Chapter 2 with regard to L1 and L2 sentence processing provided enough motivation to look into the integration of information from multiple domains (Clahsen & Felser, 2006; Ferreira, 2003; Ferreira, Bailey, & Ferraro, 2002; Ferreira, Engelhardt, & Jones, 2009; Ferreira & Patson, 2007; O’Grady, 2015, 2016; Phillips & Ehrenhofer, 2015; Sorace, 2011; Sorace & Filliaci, 2006). Linguistic knowledge is not always thoroughly attended to when a sentence can be interpreted without thoroughly parsing grammatical details even in L1 (Ferreira et al., 2002; Ferreira & Patson, 2007), and Clahsen and Felser (2006) argue that L2 speakers’ structure parsing is shallow because they rely on meaning-based information such as lexical semantics, pragmatics, and real-world knowledge.

To be able to see the effects of grammatical knowledge and real-world knowledge in action, one should ensure that both types of knowledge are available in all groups of speakers

tested. As described in the beginning of this chapter, it was already confirmed that L1 and advanced L2 speakers share the same grammatical knowledge of English articles. Below are the results from the online norming survey that tested real-world knowledge in L1 speakers and L2 speakers of various proficiency levels (5.3) and the integration experiment that pitted linguistic information against non-linguistic information (5.4).

5.3 Experiment 3a: World Knowledge Norming

For the types of non-linguistic information that generates expectations in sentence processing, associations between people (or animals) in the center and objects in the corners were used as in Figures 5.3 and 5.4. A doctor can use both stethoscopes and laptops; however, the association is stronger with a stethoscope than with a laptop (Figure 5.3). Likewise, a scientist might need a sweater as well as a microscope but a scientist is more strongly related to a microscope than to a sweater (Figure 5.4). Forty items illustrating such relationships were created. The visual stimuli were not accompanied by any spoken or written sentences. This was to ensure that no linguistic information is involved in testing participants' world knowledge.

Forty visual items were presented via a web-based experiment. The link to the web-based experiment was disseminated to L1 participants from the U.S. and L2 participants from Korea and in the U.S. In addition, a questionnaire on language learning background collected information on participants' native languages and self-reported L2 proficiency. The task was to choose a number that corresponds to the item they believe is most closely related to the person (or animal) in the center of the screen. Out of 40 items in total, the items for which participants

showed the highest agreement on the relationship (86–100%) were selected. There was no significant difference in making such associations by L1 group or L2 proficiency level.

5.4 Experiment 3b: Integration

With the confirmation that both grammatical knowledge (GK) on English articles and real-world knowledge (WK) were equally shared between L1 and all L2 speakers, the integration of the two types of knowledge was tested via a Referent Identification Task, in which participants' reaction time was measured when they identified a referent indicated by the last word of an auditory stimulus in a visual scene. A language background questionnaire and a C-test for proficiency measurement were part of the procedure.

5.4.1 Method

5.4.1.1 Participants and proficiency

Participants were recruited from the communities of University of Hawai'i at Mānoa in the U.S. and Seoul National University in Korea. Either course credit or monetary compensation was provided. Participants were asked to fill out a language learning background questionnaire via a web-based survey before they came to the laboratory. This was to ensure both L1 and L2 participants were born and raised in a monolingual environment up to age seven. The same C-test, used in Chapters 3 and 4, was used to measure L2 participants' English proficiency. Table 5.1 summarizes the demographics for both L1 and L2 participants.

Table 5.1. *Participant demographics*

		Age (years)	C-test (out of 40)	LOR (years)
L1 (n=36)	<i>Mean</i>	23.3	37.63	
	<i>SD</i>	4.76	2.0	N/A
	<i>Range</i>	18–37	33–40	
Advanced L2 (n=43)	<i>Mean</i>	23.5	35.9	0.6
	<i>SD</i>	2.30	1.5	1.78
	<i>Range</i>	18–28	34–40	0–10.5
Intermediate L2 (n=24)	<i>Mean</i>	23.48	30.1	0.86
	<i>SD</i>	2.63	2.2	2.5
	<i>Range</i>	18–30	26–33	0–13

5.4.1.2 Stimuli

Visual stimuli. Twenty-eight visual stimuli selected from the norming survey were included for the integration experiment. In fourteen items, WK pointed to a unique referent as in Figure 5.3 and in the other fourteen items, WK pointed to a non-unique referent as in 5.4. The location of unique referents was rotated around the four corners. The three referents, either in the unique condition or in the non-unique condition, were not completely identical because they needed to be seen as independent referents. Because of this, all four referents, are in a sense, unique; however, three were similar to each other while one was distinctively different from the rest, which gave the solo referent maximal uniqueness (Ahern & Stevens, 2014), which was elaborated on in Chapter 4.

Auditory stimuli. Auditory stimuli were recorded by a female native speaker of English in a sound-proof booth and were edited using the speech sound analysis software Praat (Boersma & Weenik, 2016). All auditory stimuli had a 500ms pause spliced in critical NPs between the

determiner and the noun¹⁸. The pause was placed within the critical NP of the sentence, between the article and the head noun as in (5.5) and (5.6). This was to ensure that both L1 and L2 speakers had some time to process linguistic information before they hear the head noun.

|←500ms→|

|← Phase 1 →|← Phase 2 →|← Phase 3 →|

- (5.5) a. The man will want to use *the*..... stethoscope
 b. The man will want to use *the* laptop
 c. The man will want to use *a* stethoscope
 d. The man will want to use *a* laptop

|←500ms→|

|← Phase 1 →|← Phase 2 →|←Phase 3→|

- (5.6) a. She will want to get *a* microscope
 b. She will want to get *a* sweater
 c. She will want to get *the*..... microscope
 d. She will want to get *the*..... sweater

As was seen in Chapter 3, L2 speakers are significantly slower at sentence processing and detection of grammatical anomalies than L1 speakers. If no time is given between the article and

¹⁸ The auditory stimuli were recorded with a pause between the article and the noun in the first place and the length of the pause was later modulated to be exact 500ms. This was to ensure that the vowel quality of the articles will not change under the influence of the following consonant, which can hint at what words will follow.

the head noun, L2 speakers' slower reaction might not be detected at all. By giving a 0.5-second pause to process linguistic information from *the* or *a*, the study increased the probability of observing differences in L2 speakers' behavior in comparison to L1 speakers.

The combination of the visual stimuli in Figures 5.3 and 5.4 and the auditory stimuli in (5.5) and (5.6) creates phases of information integration. At Phase 1, participants are given the visual stimuli with no sound for 2.5 seconds. After the 2.5-second preview, participants hear the auditory stimulus. Phase 1 extends from the beginning of the preview time until the end of the verb. Participants will form their initial expectation based on their world knowledge at this point. After a 0.5-second pause, the crucial linguistic information *the* or *a* is given and this is Phase 2. At this stage, participants will judge if the linguistic information aligns or misaligns with the initial expectation they had at Phase 1. In the case of alignment, the initial expectation will be reinforced but it might be reconsidered in the case of misalignment. Finally at Phase 3, participants hear the last word of the sentence, at which point they will be able to confirm whether or not the initial expectation they had matches the actual last word of the sentence. If this proposition of incremental information integration is correct, information presented at each phase will generate expectations or revise previous expectations.

Fillers. A total of fifty-six fillers (Appendix 3b) were added to disguise the research purpose. Seven fillers had one plausible option each and another seven had three plausible options. Another twenty-one fillers had two plausible options as in Figure 5.5, and the last twenty-one fillers had all four options as plausible ones, as illustrated in Figure 5.6. The number of plausible options was manipulated this way so that participants would think that the trials are grouped into four categories of equal numbers of items with one to four plausible options. This was intended to obfuscate the patterns of the critical stimuli, hence, the research purpose.

|←500ms→|

|← Phase 1 →|← Phase 2 →|←Phase 3→|

(5.7) The woman will want to bake some desserts.

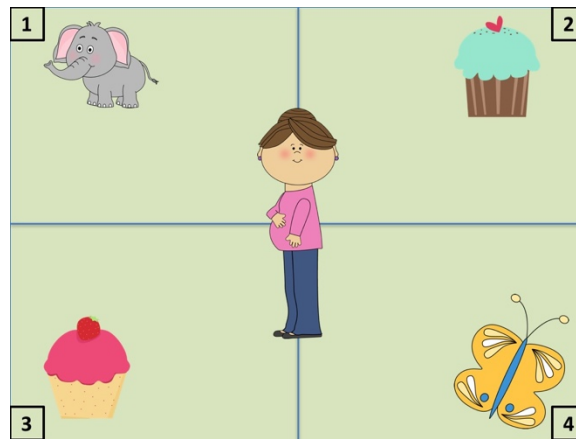


Figure 5.5. A visual stimulus example for fillers with two plausible options.

|← 500ms→|

|← Phase 1 →|← Phase 2 →|←Phase 3→|

(5.8) The boy will want to try on a new pair of ski boots.

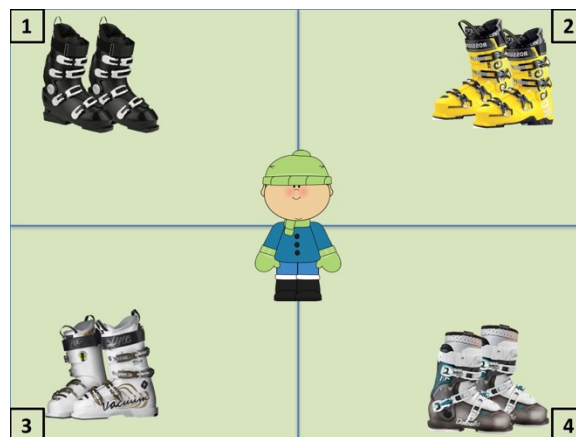


Figure 5.6. A visual stimulus example for fillers with four plausible options.

For filler items, the pause was placed either right before the last NP of the sentence if the last word of the sentence was a mass noun (e.g., “The baby will want to drink... milk”). When other determiners (5.7) or classifiers (5.8) were used, the pause was placed between a determiner and the head noun. No filler items had an article within the last NP to ensure that participants’ reaction to articles in critical trials are not affected by the use of English articles in fillers.

5.4.1.3 Procedure

Once they arrived at the laboratory, participants were seated in a sound attenuated booth. Recorded audiovisual instructions for the experimental tasks were given on a computer screen, followed by practice trials and the main experiment. The task was to identify the referent indicated by the last word of the sentence as quickly as possible. Participants were told to wait until they hear the last word of the sentence and press a button with a number that corresponds to the referent they believe is denoted by the last word of the sentence. Unlike in Experiment 2 in Chapter 4, beep sounds were unnecessary. The last word was a cue for participants to press a button and the trial ended at the time of their button press or 5 seconds after the sentence ended. The entire procedure took about 17 minutes for each participant.

5.4.2 Analysis

Reaction time was measured between the offset of the final word of the sentence to the button press in reaction to the auditory stimuli. Since items had different numbers of plausible options, residual analysis was unavoidable. That is, when there is only one plausible option,

reaction time will be faster than when there are more than one. Therefore, comparing raw RTs across items with different numbers of plausible options could result in creating artifacts that are not related to factors under investigation. Following is a detailed description of statistical analysis of reaction times.

5.4.2.1 Log Transformation.

As was in Chapter 3, the raw reaction time was transformed for logarithms and fit in a regression model to calculate residuals. This was to satisfy the assumptions of linear regression: the linearity of dependent and independent variables and the normality of errors. First, Table 5.2 shows that the skewness and kurtosis values were improved in log-transformed RT data. The raw and log-transformed RTs were compared by group in an lmer model (5.9).

Table 5.2. *Descriptive statistics for reaction time (RT) data before and after log transformation.*

data	group	n	mean	sd	min	max	skew	kurtosis	se
Raw RT	L1	2852	953.55	904.24	1	6289	2.01	5.23	16.93
	adv L2	3497	998.9	882.08	1	7013	1.73	3.73	14.92
	Inter L2	1914	1505.61	1194.06	3	7061	1.22	1.40	27.29
Log RT	L1	2852	6.42	1.05	0	8.75	-0.95	2.71	0.02
	adv L2	3497	6.5	1	0	8.86	-0.87	1.99	0.02
	Inter L2	1914	6.93	1.01	1.1	8.86	-1.14	2.65	0.02

```
(5.9) model1=lmer(RT ~ group + (1|Subject), data)
```

```
(5.10) model2=lmer(logRT ~ group + (1|Subject), data)
```

The results of pairwise comparisons between groups show that advanced L2 speakers were not significantly different from L1 speakers in either raw ($\beta=53.01$, t -value= 0.41, $p=0.68$) or log-transformed ($\beta=0.12$, t -value= 0.91, $p=0.37$) RTs but intermediate L2 speakers showed a significant difference in both raw and log-transformed values compared to both L1 (raw: $\beta=514.82$, t -value=3.94, $p=0.000$; log: $\beta=0.55$, t -value= 4.17, $p=0.000$) and advanced L2 speakers (raw: $\beta=461.81$, t -value= 3.96, $p=0.000$; log: $\beta=0.43$, t -value= 3.65, $p=0.000$).

5.4.2.2 Residual reaction time and linear mixed effects regression.

Next, the log-transformed RTs were put in a regression model (5.11) to address factors that influence RTs. As can be seen in Appendix 3a, the final word of each stimulus differed from item to item. Therefore, the length of critical word (in *ms*) was included in the residual regression analysis (`wordLength`). RTs also varied among items with different numbers of plausible options and they did not increase linearly as the number of plausible options increased; therefore, the number of plausible options was included in the residual analysis model as log values (`optionNumber`). For the same reason, trial order (`trialOrder`) was also put in as log-transformed values. (5.11) is the linear mixed effects regression (`lmer`) code used in R (R Core Team, 2016) to calculate the regression line with random intercepts for subject and (5.12) is the residual code that calculates the discrepancies between the expected values of the regression model (5.11) and the observed values in the data collected.

```
(5.11) model=lmer(logRT ~ wordLength + log(optionNumber) + log(trialOrder) +
  (1|Subject), data)
(5.12) residuals(model)
```

Table 5.3. *Linear Mixed Effects Regression output for log RT residual analysis in L1 data*

	β (Estimate)	Std. Error	t value	p value	
(Intercept)	6.1917	0.1263	48.9902	0.0000	***
length of critical word	0.0000	0.0000	-1.1288	0.2589	
log(# of plausible options)	0.6196	0.0302	20.4569	0.0000	***
log(trial order)	-0.0692	0.0167	-4.1344	0.000	***
· $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$					

Table 5.4. *Linear Mixed Effects Regression output for log RT residual analysis in adv. L2 data*

	β (Estimate)	Std. Error	t value	p value	
(Intercept)	6.5872	0.1046	62.9477	0.0000	***
length of critical word	0.0001	0.0000	-2.3260	0.0200	**
log(# of plausible options)	0.4697	0.0236	19.8978	0.0000	***
log(trial order)	-0.1040	0.0132	-7.8995	0.0000	***
· $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$					

Table 5.5. *Linear Mixed Effects Regression output for log RT residual analysis in inter. L2 data*

	β (Estimate)	Std. Error	t value	p value	
(Intercept)	7.0681	0.1382	51.1569	0.0000	***
length of critical word	0.0000	0.0001	-0.9725	0.3307	
log(# of plausible options)	0.7557	0.0338	22.3164	0.0000	***
log(trial order)	-0.1994	0.0191	-10.4399	0.0000	***
· $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$					

Tables 5.3 to 5.5 show that all participants reacted more slowly when there were more options to choose from (# of plausible options) and faster when they became familiar with the task towards the end of the experiment (trial order). However, the length of critical words (length of critical word) did not have a significant effect on L1 speakers (Table 5.3) or intermediate L2 speakers (Table 5.5); however, as shown in Table 5.4, there was an effect on the

advanced L2 speakers (Table 5.4). To make comparison easier amongst the three groups, the length of the critical word was left in the regression models of all three groups.

With the regression models for each group (5.11), residuals were calculated by measuring the deviations of each data point from the fitted model (5.12). Linear mixed effects regression models were fitted with the residual values for each group (5.13). As was in Chapter 4, lmer models were selected using Akaike Information Criterion and Bayesian Information Criterion when a full model with both slopes and intercepts for random effects did not merge. The lmer model for L1 speakers (5.13a) include random intercepts for both subjects and items but slopes were included only for world knowledge (WK) and not for grammatical knowledge (GK). (5.15b) is a mathematical notation for the regression model (Raudenbush & Bryk, 2002).

For L2 speakers, lmer models were fitted with random intercepts for both subjects and items but their slopes were included only for GK. The lmer model is given in (5.13c) and its mathematical notation is left out for want of space.

(5.13a) `lmer(residual~GK*WK+(1+WK|Subject)+(1+WK|item), data)`

GK: Grammatical knowledge (match vs. mismatch at Phase 2); +G OR -G

WK: World knowledge (match vs. mismatch at Phase 3); +W or -W

$$(5.13b) Y_{si} = \beta_{00} + \gamma_{0s} + \delta_{0i} + \beta_1 * X + (\beta_2 + \gamma_{2s} + \delta_{2i}) * W + \beta_3 * XW + e_{si}$$

Y_{si} is the log-transformed reaction time residual for subject s in item i ;

X is the fixed factor ‘grammatical knowledge (GK)’;

W is the fixed factor ‘world knowledge (WK)’;

β_{00} is the grand mean of the log-transformed reaction time residual;

- γ_{0s} is the subject intercept, the deviation from β_{00} for subject s ;
- δ_{0i} is the item intercept, the deviation from β_{00} for item i ;
- β_1 is the main effect of the fixed factor X ;
- β_2 is the main effect of the fixed factor W ;
- γ_{2s} is the subject slope with regard to the fixed factor W ;
- δ_{2i} is the item slope with regard to the fixed factor W ;
- β_3 is the interaction effect of the fixed factors XW ;
- e_{si} is the observation-level error with its mean 0 and variance σ^2 .

5.4.3 Results

As was seen in examples (5.5) and (5.6) repeated below, linguistic information at Phase 2 could either align or misalign with the initial prediction at Phase 1. Therefore, the conditions where the article at Phase 2 aligned with the Phase 1 prediction were labeled [+G] and the conditions where the article misaligned with the initial prediction were labeled [−G].

Likewise, the final word of the sentence (in Phase 3) could either confirm or deny the initial prediction (at Phase 1) that was derived from world knowledge. The matching conditions were labeled [+W] and the mismatching conditions [−W]. The regression model in (5.13) included the two phases as fixed effects [G] and [W] with the two levels *match* [+] and *mismatch* [−] in each factor.

What should be noted is that [−G] does not indicate ‘ungrammaticality’ but the ‘misalignment of linguistic information with Phase 1 expectation based on non-linguistic

information. This should be carefully looked at especially in (5.5d) and (5.6d) where a laptop and the sweater are grammatical but do not match the expectation generated Phase 1.

	←	Phase 1	→ ←	Phase 2	→ ←	Phase 3	→	
(5.5)	a.	The man will want to use <i>the</i>	stethoscope	[+G		+W]		
	b.	The man will want to use <i>the</i>	laptop	[+G		–W]		
	c.	The man will want to use <i>a</i>	stethoscope	[–G		+W]		
	d.	The man will want to use <i>a</i>	laptop	[–G ¹⁹		–W]		
	←	Phase 1	→ ←	Phase 2	→ ←	Phase 3	→	
(5.6)	a.	She will want to get <i>a</i>	microscope	[+G		+W]		
	b.	She will want to get <i>a</i>	sweater	[+G		–W]		
	c.	She will want to get <i>the</i>	microscope	[–G		+W]		
	d.	She will want to get <i>the</i>	sweater	[–G		–W]		

5.4.3.1 L1 speakers

L1 speakers' reaction time (RT) changes can be seen at a glance in Figure 5.7. Residual plots should be understood as how much each condition deviates from the mean in reaction time. The Y-axis of the plot indicates the residuals of log-transformed reaction time and the horizontal

¹⁹ As is mentioned in the text, please note that the label [–G] does not indicate the articles *a* in (5.5d) and *the* in (5.6d) are ungrammatical. It indicates that the grammatical information the articles convey does not match the initially expected referent at Phase 1 (i.e., stethoscope in (5.5) and microscope in (5.6)).

line indicates the average amount of time L1 participants took to respond to the stimuli. This mean is calculated by including not only critical trials but also filler trials. Bars hanging below the horizontal line indicates that RTs for those conditions were shorter than average while the third bar, which goes above the horizontal line, shows that the third condition induced longer RTs than average.

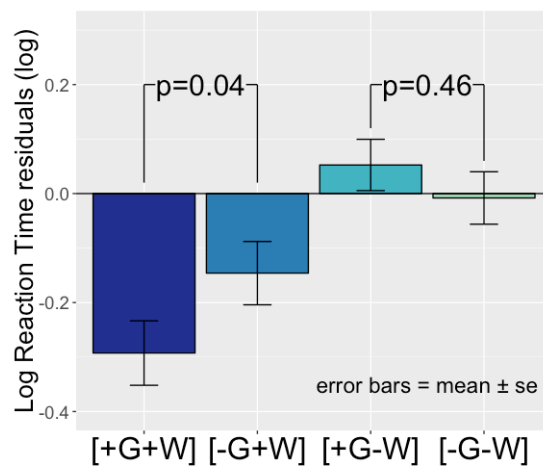


Figure 5.7. L1 speakers' log reaction time residuals by condition.

The first dark blue bar is the condition in which both linguistic information ([+G]) and non-linguistic information ([+W]) matched their initial expectation. This condition led to the shortest reaction time among the four conditions. When linguistic information didn't align with their initial expectation ([-G]) but non-linguistic information did ([+W]), the reaction time goes up significantly compared to the first condition. The same goes for the condition in which linguistic information aligned ([+G]) but non-linguistic information misaligned ([-W]) with the initial expectation, which is indicated by the third bar.

At this point, one can see that misalignment between the initial expectation and the information available at later phases can lead to significant increases in reaction time. With this

logic, one might expect to see double the amount of surprise at the fourth condition $[-G, -W]$, where the initial expectation is supported at neither of the two later phases; however, the fourth bar in Figure 5.7 is shorter than the third bar. This decrease in reaction time in the fourth condition in comparison to the third condition, I argue, is important evidence that L1 speakers of English incrementally process information generating initial expectations and revising them based on incoming information at different stages.

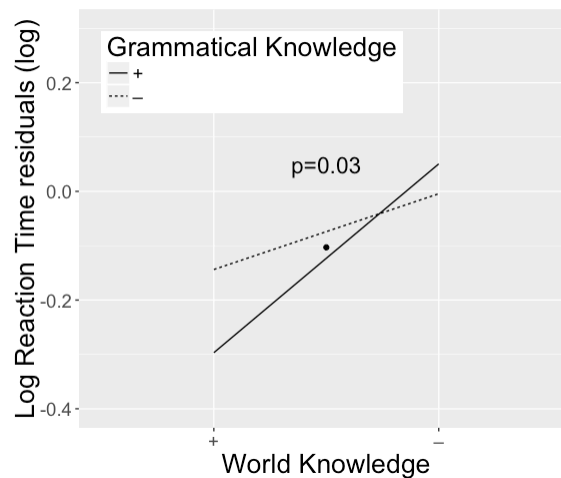


Figure 5.8. Interaction of grammatical knowledge and world knowledge in L1 speakers.

One might think that the decrease in RTs from the third condition $[+G, -W]$ to the fourth condition $[-G, -W]$ is minimal and that the fourth condition induces still a significantly higher RT compared to the first and second conditions. However, Figure 5.8 shows how grammatical knowledge and world knowledge interact with each other and linguistic information that aligns with the initial expectation ($[+G]$) reduces RT when non-linguistic information also aligns with the initial expectation ($[+W]$) but has the opposite effect in the other case ($[-W]$). That is, the solid line ($[+G]$) shows a lower RT than the dotted line ($[-G]$) in the $[+W]$ condition but the

opposite pattern in the [-W] condition. This interaction was significant ($p=.03$) and certainly indicates that L1 speakers generate their expectation at Phase 1 and revise it at Phase 2 if the linguistic information does not align with Phase 1. In other words, they process both linguistic and non-linguistic information incrementally.

Table 5.6. *Main and interaction effects of Grammatical Knowledge and World Knowledge in L1.*

(a) Fixed effects

		(Estimate)	Std. Error	t value	p value	
(Intercept)	β_0	-0.1032	0.0515	-2.0030	0.045	*
Phase 2 (GK)	β_1	-0.0239	0.0250	-0.9558	0.339	
Phase 3 (WK)	β_2	-0.1175	0.0424	-2.7672	0.006	**
Phase 2 x Phase 3	β_3	-0.0530	0.0250	-2.1150	0.034	*

* $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$

(b) Random effects

Groups	Name	Variance	Std.Dev.	Corr
Subject	(Intercept)	γ_{0s}	0.0087	0.0934
	Phase 3	γ_{2s}	0.0112	0.1056
item	(Intercept)	δ_{0i}	0.0501	0.2239
	Phase 3	δ_{2i}	0.0243	0.1560
Residual			0.5672	0.7531

Number of obs: 923, groups: Subject, 36; item, 28

Table 5.6a summarizes the main effects of grammatical knowledge (GK) and world knowledge (WK) and the interaction of the two and Table 5.6b shows the random effects of subjects and items. The intercept in Table 5.6a is the mean of all four conditions ($\beta_0=-0.103$), which is indicated as the dot in the center in Figure 5.8. As discussed earlier, the horizontal line in Figure 5.7 is the mean RT for all trials in the experiment. The intercept in Table 5.6 shows the mean RT for only the critical trials. It shows that the critical trials took less time than the average of all trials including critical and filler trials. Linguistic information that becomes available at

Phase 2 ($\beta_1=-0.024$) does not make a significant difference as is indicated on the second row. On the other hand, non-linguistic information that becomes available at Phase 3 makes a significant difference. Reaction time goes lower than the intercept ($\beta_2=-0.118$) in the world knowledge (WK)-matching condition (+W) and higher in the WK-mismatching condition (-W). Finally, the interaction row indicates that when the feature values of GK and WK are the same (e.g., both + or both -), the reaction time will decrease by the interaction estimate ($\beta_3=-0.053$).

Using the estimates (β) in Table 5.6a, log RT residuals can be calculated for each condition. For the [+G, +W] condition, all four estimate values can be added up: $\beta_0 + \beta_1 + \beta_2 + \beta_3 = (-0.1032) + (-0.0239) + (-0.1175) + (-0.0530) = -0.2976$, which is close to the observed value of -0.297 in Table 5.7.

Table 5.7. *The log RT residuals by condition in L1.*

Conditions	N	log RT residual	S.D.
[+G+W]	228	-0.297	1.019
[-G+W]	230	-0.144	1.000
[+G-W]	234	0.051	0.796
[-G-W]	231	-0.004	0.821

The same calculation can be applied to the [-G, -W] condition. As was explained earlier, mismatching linguistic and non-linguistic information will result in the increase of RT estimates. And for interaction, as long as the two factors have the same feature values (both - in this case), a decrease in RT is expected: $\beta_0 - \beta_1 - \beta_2 + \beta_3 = (-0.1032) - (-0.0239) - (-0.1175) + (-0.0530) = -0.0148$. The calculated values from the estimates in Table 5.6a do not show a 100% correspondence to the observed values in Table 5.7 due to the random variances (Table 5.6b). However, the estimates of fixed effects provide a picture of how reading times are affected by linguistic and

non-linguistic information and their interaction. Without the interaction taken into account, the [–G, –W] condition should have shown a much longer RT residual (0.0382), but the significant interaction effect lowered the log RT residuals to –0.004 in Table 5.7—although not as low as –0.0148 as was estimated by the lmer model output in Table 5.6a.

Table 5.8. *Main and interaction effects of GK, WK, and article in L1.*

	β (Estimate)	Std. Error	t value	p value	
(Intercept)	–0.102	0.052	–1.963	0.050	.
Phase 2 (GK)	–0.023	0.025	–0.928	0.354	
Phase 3 (WK)	–0.117	0.042	–2.798	0.005	**
article (a/the)	0.016	0.025	0.647	0.518	
Phase 2:Phase 3	–0.053	0.025	–2.121	0.034	*
Phase 2:article	–0.025	0.050	–0.493	0.622	
Phase 3:article	–0.011	0.025	–0.432	0.666	
Phase 2:Phase 3:article	0.052	0.038	1.355	0.175	

. $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$

Before moving on to advanced L2 speakers, it needed to be confirmed that there was no difference between conditions where a definite article was felicitous and those where an indefinite article was felicitous. Table 5.8 shows the results of an lmer analysis where GK, WK, and article were factored into as fixed effects. The output is not different from when article was not included in the analysis (Table 5.7).

A graphic illustration is provided in Figure 5.9. The first four bars show the conditions in which WK predicted a unique referent (e.g., doctor—one stethoscope; unique-referent condition) and the last four show the conditions in which WK predicted unique referents (e.g., scientist—three microscopes; non-unique referent condition). The general pattern is the same between the

unique and non-unique referent conditions in that RT increases from the first condition to the third and it slightly decreases from the third condition to the fourth.

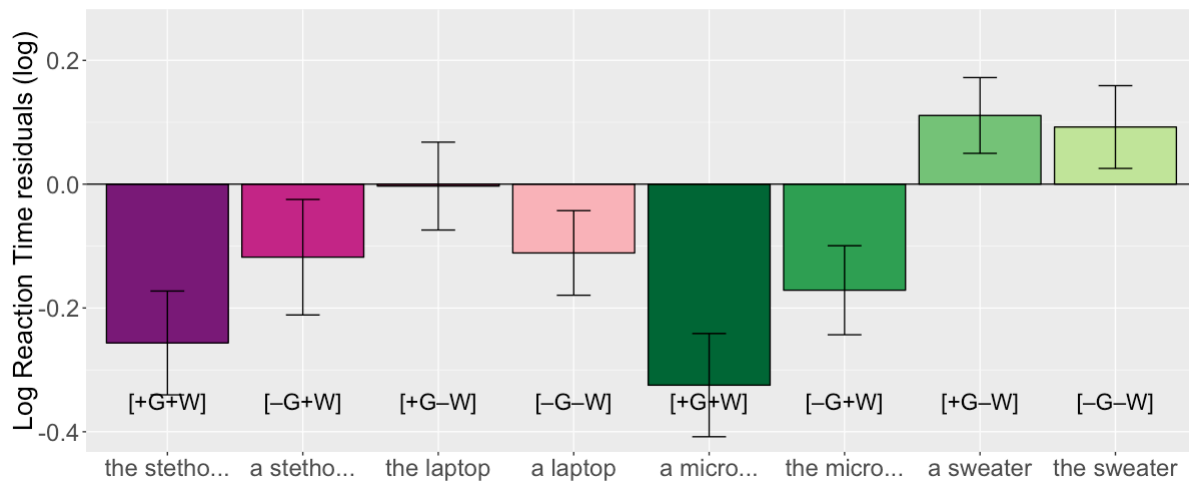


Figure 5.9. L1 speakers' log reaction time residuals by condition and visual stimuli type

This subtle, if not significant, decrease from the third to the fourth bar is the indication that the $[-G]$ linguistic information is taken into account at Phase 2 to revise their initial prediction from Phase 1. If no interaction had been in effect, the fourth bar should have shown a longer RT than the $[+G, -W]$ condition. Even if the pairwise comparison between the $[+G, -W]$ and the $[-G, -W]$ conditions was not significant, the fact that the $[-G, -W]$ was not higher than $[+G, -W]$ implies that L1 speakers at Phase 2 re-evaluates their prediction from Phase 1.

In sum, there was no main effect of linguistic information ($[\pm G]$) and a significant main effect of non-linguistic information ($[\pm W]$). Also, the interaction effect of linguistic and non-linguistic information was also significant. The results indicate that L1 speakers process both linguistic and non-linguistic information incrementally, which means that they generate and revise their expectations bit by bit based on information as it becomes available.

5.4.3.2 Advanced L2 speakers

Figure 5.10 provides a quick glance at the results of advanced L2 data analysis. The pattern of RT increase/decrease is somewhat different from that of L1 speakers. What is most noteworthy is that the RT difference between the first and second conditions is not significant but that between the third and fourth conditions is significant. This is the exact opposite pattern from what we have just seen in L1 speakers.

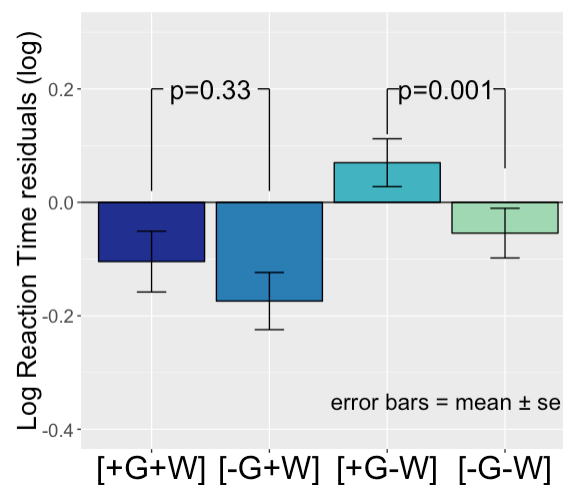


Figure 5.10. Advanced L2 speakers' log reaction time residuals by condition.

This difference indicates that the (mis)alignment of grammatical knowledge ($[\pm G]$) did not make a difference when non-linguistic information alone could predict upcoming linguistic material ($[+W]$). $[\pm G]$ could make a difference only when upcoming linguistic material could not be predicted based on non-linguistic information ($[-W]$). It is true that $[-W]$ sentences led to higher RTs in general. However, the (mis)alignment of linguistic information did make an important difference in either $[+W]$ or $[-W]$ conditions in L1 speakers. Also, the cross-over

interaction was observed through the RT decrease from the third condition to the fourth, which attested to the incremental processing of linguistic and non-linguistic information. This pattern was not observed in advanced L2. Table 5.9 shows that there were significant main effects of both linguistic (GK) and non-linguistic information (WK), but no interaction effect.

Table 5.9. *Main and interaction effects of GK and WK in advanced L2 speakers.*

(a) Fixed effects

	β (Estimate)	Std. Error	t value	p value	
(Intercept)	-0.0627	0.0583	-1.0773	0.2814	
Phase 2 (GK)	0.0452	0.0198	2.2793	0.0226	*
Phase 3 (WK)	-0.0745	0.0193	-3.8619	0.0001	***
Phase 2:Phase 3	-0.0179	0.0194	-0.9245	0.3552	

· $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$

(b) Random effects

Groups	Name	Variance	Std.Dev.	Corr
Subject	(Intercept)	γ_{0s} 0.0069	0.0829	
	Phase 2	γ_{1s} 0.0007	0.0259	-1.0000
item	(Intercept)	δ_{0i} 0.0802	0.2832	
Residual		0.4239	0.6511	

Number of obs: 1145, groups: Subject, 43; item, 28

Figure 5.11 shows that advanced L2 speakers showed higher RTs in the [+G] conditions than in the [-G] conditions. The RT increase from [-G] to [+G] was maintained in both [+W] and [-W] conditions. This could be misinterpreted as evidence that advanced L2 speakers have completely opposite grammatical knowledge in comparison to L1 speakers. The behavior of advanced L2 speakers, however, can be better understood when the four conditions are broken down into the unique and non-unique referent conditions as in Figure 5.12.

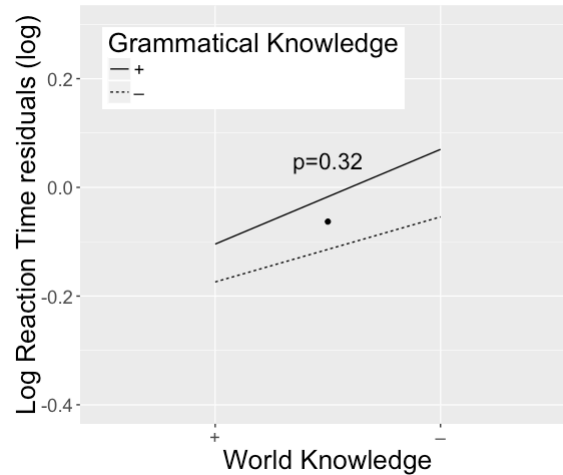


Figure 5.11. Interaction of grammatical knowledge and world knowledge in advanced L2.

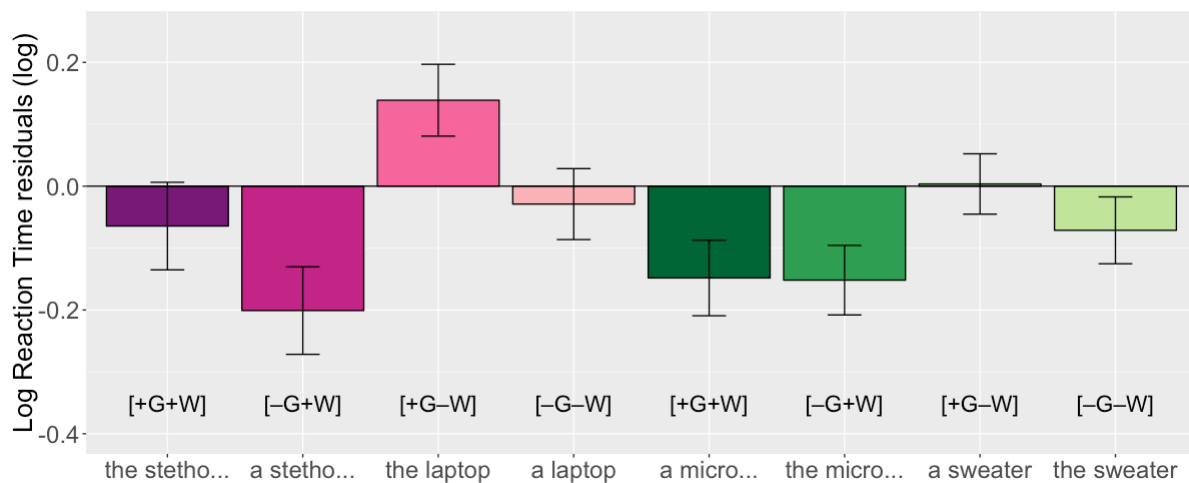


Figure 5.12. Log reaction time residuals by condition and visual stimuli type in advanced L2

Comparing Figures 5.9 and 5.12 makes it clear that L1 and advanced L2 speakers behave differently in processing the two sources of information. L1 speakers show the same pattern of RT increase/decrease in both unique and non-unique referent conditions (Figure 5.9) but advanced L2 speakers show different patterns of RT increase/decrease between the unique and non-unique referent conditions (Figure 5.12). The pattern across the four conditions in the unique

referent condition (the first four bars in pink) is different from that in the non-unique referent condition (the latter four bars in green). In the unique referent condition, *a stethoscope* [-G, +W] induced shorter RTs than *the stethoscope* [+G, +W]. This could be interpreted as advanced L2 speakers' interpreting the indefinite article as a singularity marker if it were not for the pattern observed in the non-unique referent condition (*a sweater* vs. *the sweater*). To argue that advanced L2 speakers use *a/an* for singularity marking, *a sweater* should induce shorter RTs than *the sweater* since, in the non-unique referent condition, 'sweater' is the item that stands alone. However, that is not the case. Moreover, Experiment 2 in Chapter 4 showed that advanced L2 speakers could use English articles in a native-like manner; therefore, it does not make sense to conclude that advanced L2 speakers use indefinite articles as a singularity marker.

Next, let's look at the conditions in which the final word of the sentence indicated a non-unique referent (laptop vs. microscope). Definite and indefinite articles do not make a difference when real-world knowledge alone can predict the last word of the sentence. That is, *the microscope* and *a microscope* is not meaningfully differentiated in reaction time. In contrast, articles make a significant difference in reaction to *the laptop* and *a laptop*. Although there were three options to choose from when either laptop or microscope is the last word of the sentence, the choice of article makes a significant difference only when the last word is not what real-world knowledge predicted at Phase 1.

To account for both cases of definite and indefinite articles, it should be determined whether [\pm G] can be processed as quickly as [+W] and whether real-world knowledge alone predicts upcoming linguistic material. In general, reaction time was significantly slower when the last word of the sentence was not the one predicted by WK [-W] than when the sentence-final word was [+W]. However, by the time [-W] information was processed, [\pm G] processing

was also available, which led to significant differences in RTs depending on the (mis)alignment of linguistic information with initial predictions. This difference is important in that it means advanced L2 speakers can make active use of linguistic information although their processing of linguistic information is much slower than L1 speakers. This means that compared to L1 speakers who were sensitive to $[\pm G]$ information at the same time they reacted faster to $[+W]$ information, advanced L2 speakers could react just as fast to $[+W]$ information as L1 speakers but their reaction to $[\pm G]$ was delayed as late as they could react to $[-W]$ information. This timing difference between L1 and advanced L2 speakers will be discussed in terms of the imbalanced automaticity of linguistic and non-linguistic information in L2 sentence processing in the discussion section of this chapter.

Before reporting the analysis results of the intermediate L2 speaker group, the log RT residuals of advanced L2 speakers and an lmer model with article as an added fixed factor are reported in Table 5.10 and 5.11, respectively. Table 5.10 summarizes log RT residuals in each condition and Table 5.11 shows that article itself did not make a difference. There was no main effect of article or interaction of article and other fixed effects (GK and WK). This means that, as in L1 speakers, definite and indefinite articles per se will not interact with either GK or WK.

Table 5.10. *The log RT residuals by condition in advanced L2.*

Conditions	<i>N</i>	log RT residual	<i>S.D.</i>
[+G+W]	282	-0.104	0.901
[-G+W]	278	-0.174	0.840
[+G-W]	297	0.070	0.727
[-G-W]	288	-0.054	0.743

Table 5.11. *Main and interaction effects of GK, WK, and article in advanced L2 speakers.*

	β (Estimate)	Std. Error	t value	p value	
(Intercept)	-0.062	0.059	-1.052	0.293	
Phase 2 (GK)	0.046	0.019	2.344	0.019	*
Phase 3 (WK)	-0.074	0.019	-3.810	0.000	***
article (a/the)	0.031	0.019	1.579	0.114	
Phase 2:Phase 3	-0.017	0.019	-0.889	0.374	
Phase 2:article	0.020	0.058	0.343	0.731	
Phase 3:article	-0.005	0.019	-0.282	0.778	
Phase 2:Phase 3:article	0.022	0.019	1.149	0.250	

· $p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$

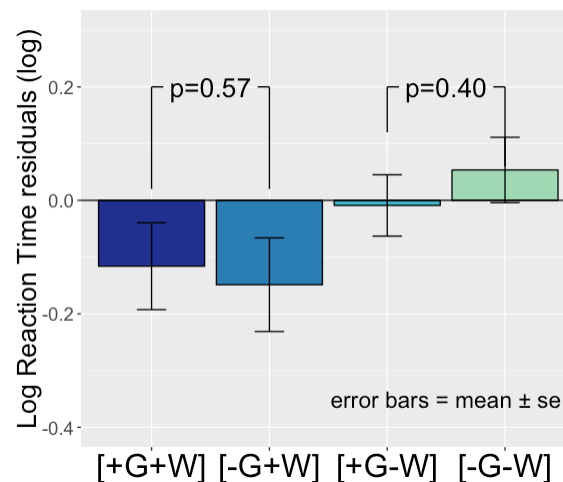


Figure 5.13. Intermediate L2 speakers' log reaction time residuals by condition.

5.4.3.3 Intermediate L2 speakers

Finally, let us consider the intermediate L2 group. They seemingly display similar behavior to that of advanced L2 speakers in the [+W] conditions but different behavior from that of both L1 and advanced L2 speakers in the [-W] conditions. Figure 5.13 shows that the first two bars on the plot have the same pattern as for advanced L2 speakers, with the [+G] condition

showing longer RT than the [-G] condition. In contrast, the latter two bars on the plot show a different pattern from both L1 and advanced L2 groups. This apparent pattern of behavior in the intermediate group needs to be examined in depth.

First of all, the intermediate group showed only the main effect of world knowledge (Table 5.12). Table 5.13 summarizes log RT residuals in each condition in the intermediate L2 group and Table 5.14 shows that even when article is factored into analysis, no effects are observed other than the main effect of world knowledge. With only a small sample size, the data from the intermediate L2 group may seem to have no meaningful implications; however, one should remember two things about intermediate L2 speakers' behavior with regard to English articles: (i) their interpretation of definite and indefinite articles from Chapter 4 and (ii) the reaction time difference amongst the three groups in the current experiment.

Table 5.12. *Main and interaction effects of GK and WK in intermediate L2 speakers.*

(a) Fixed effects

	β (Estimate)	Std. Error	t value	p value	
(Intercept)	-0.0467	0.0565	-0.8224	0.4094	
Phase 2 (GK)	-0.0109	0.0289	-0.3792	0.7045	
Phase 3 (WK)	-0.0863	0.0288	-2.9952	0.0027	**
Phase 2:Phase 3	0.0341	0.0288	1.1845	0.2362	

$p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$

(b) Random effects

Groups	Name	Variance	Std.Dev.	Corr
Subject	(Intercept)	γ_{0s} 0.0000	0.0000	
	Phase 2	γ_{1s} 0.0000	0.0000	1.0000
item	(Intercept)	δ_{0i} 0.0665	0.2578	
Residual		0.5175	0.7193	

Number of obs: 632, groups: item, 28; Subject, 24

Table 5.13. *The log RT residuals by condition in intermediate L2*

Conditions	<i>N</i>	log RT residual	<i>S.D.</i>
[+G+W]	152	-0.116	0.944
[-G+W]	164	-0.149	1.057
[+G-W]	159	-0.009	0.682
[-G-W]	157	0.054	0.722

Table 5.14. *Main and interaction effects of GK, WK, and article in intermediate L2 speakers.*

	β (Estimate)	Std. Error	t value	p value	
(Intercept)	-0.047	0.057	-0.819	0.413	
Phase 2 (GK)	-0.011	0.029	-0.384	0.701	
Phase 3 (WK)	-0.085	0.029	-2.956	0.003	**
article	0.000	0.029	0.006	0.996	
Phase 2:Phase 3	0.035	0.029	1.204	0.229	
Phase 2:article	0.042	0.057	0.746	0.456	
Phase 3:article	0.018	0.029	0.620	0.535	
Phase 2:Phase 3:article	0.029	0.029	0.997	0.319	

$p < .1$; * $p < .05$; ** $p < .01$; *** $p < .001$

Making a direct comparison between the intermediate group (Figure 5.14) and the L1 (Figure 5.9) and advanced L2 (5.12) groups might give the false impression that intermediate learners behave similarly to advanced learners in the [+W] conditions and differently in the [-W] conditions. However, in Chapter 4, intermediate L2 learners predicted unique referents numerically more often when an indefinite article is given. One might therefore speculate that the intermediate learners in this experiment might have been doing the same thing. To test this, I inversed the grammaticality feature [\pm G] in Figure 5.15. That is, for intermediate L2 speakers, the indefinite article *a* could be [+G] when the visual stimulus had the doctor in the center and one stethoscope and three laptops in the corners. Likewise, the indefinite article *a* could be [-G]

if the visual stimulus had the scientist in the center and three microscopes and a sweater in the corners.

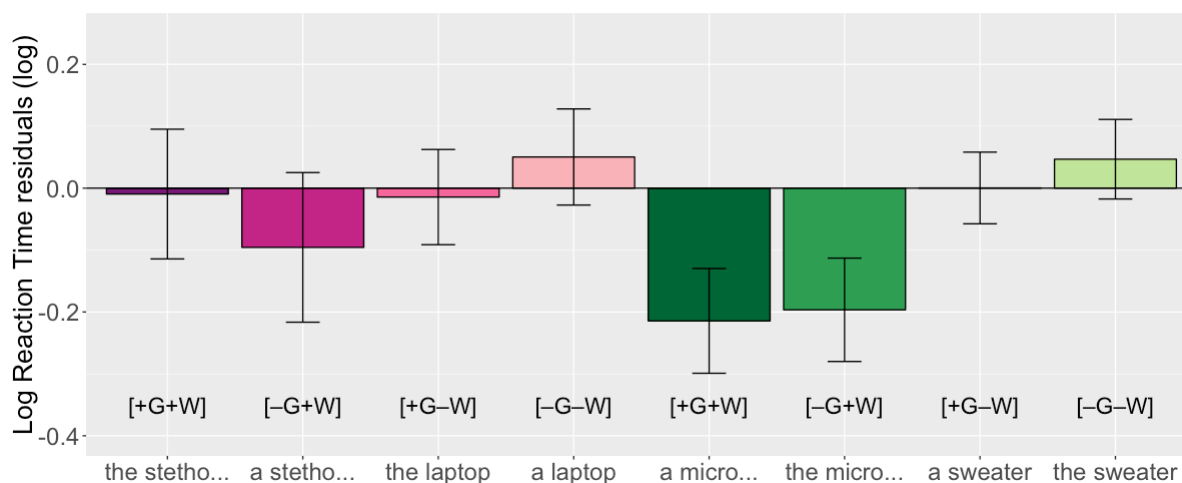


Figure 5.14. Log reaction time residuals by condition and visual stimuli type in intermediate L2

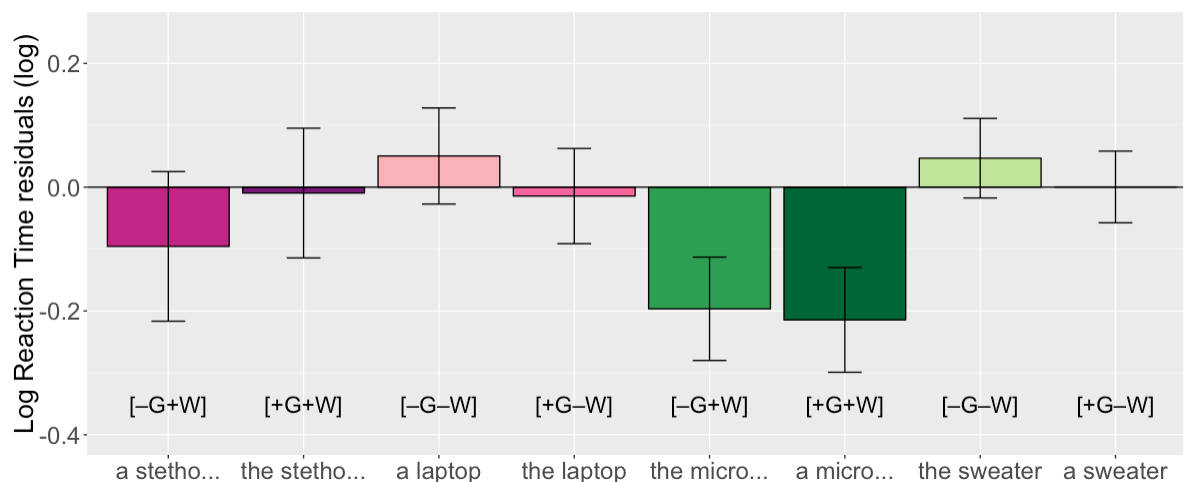


Figure 5.15. Intermediate L2 log RT residuals plot with inversed grammaticality

What Figure 5.15 shows, then, is the possibility that intermediate L2 speakers could be using both linguistic and non-linguistic information incrementally despite their non-targetlike

Interlanguage grammar (Selinker, 1972). That is, in the visual stimulus of Figure 5.3, where a doctor and a stethoscope are associated, an intermediate L2 speaker can react faster when an indefinite article is given at Phase 2 than when a definite article is given. For the same reason, they can react to *a sweater* faster than *the sweater* in the visual stimulus of Figure 5.4, where a scientist and microscopes are associated. In Figures 5.3 and 5.4, the unique referents correspond to the stethoscope and the sweater, respectively. When an article information is provided at Phase 2, participants can use this information to either confirm or revise their initial expectations from Phase 1 when reacting to the last word of the sentence at Phase 3.

The same logic applies to non-unique referents as well. Unlike L1 and advanced L2 speakers who react faster to *a laptop* than to *the laptop*, intermediate L2 speakers find it hard to process *a laptop* since there are three laptops in Figure 5.3. When real-world knowledge predicts non-unique referents, as in Figure 5.4, an indefinite article is not particularly useful and the contrast between *a microscope* and *the microscope* might not be part of their Interlanguage grammar given intermediate L2 speakers' behavior in Chapter 4. This interpretation of the intermediate L2 data naturally leads to the question of why their behavior is not viewed in terms of whether real-world knowledge alone can predict upcoming linguistic material. Earlier, I argued that, when world knowledge alone can generate predictions, advanced L2 speakers focus their attention on world knowledge information.

This is the point at which the slower reaction time the intermediate group displayed comes into play. The analysis section (5.4.2.2) reported the comparison of raw and log RTs amongst the three groups (Table 5.2), and the intermediate group showed longer RTs compared to both L1 and advanced L2 groups. That is, intermediate L2 speakers took a longer time to respond to the auditory and visual stimuli. This could mean that the extra time they took allowed

them to integrate linguistic information even when real-world knowledge alone could predict upcoming linguistic material. The pattern of cross-over interaction can be observed in Figure 5.16.

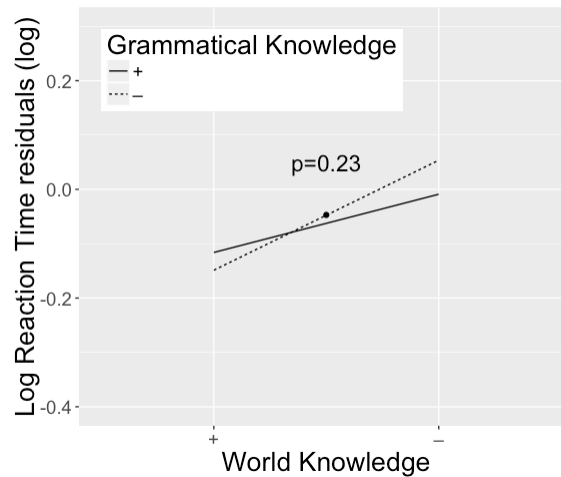


Figure 5.16. Interaction of grammatical knowledge and world knowledge in intermediate L2.

As Tables 5.12 and 5.14 showed, the only significant effect in the intermediate group involved world knowledge. It may be odd that such an elaborate interpretation of this fact is provided here. The basis for this elaboration lies in possible availability of a large sample size for intermediate L2 speakers. There were only 24 participants in the intermediate group and the significant interaction was not observed, unlike in L1 and advanced L2 speakers; however, the pattern seen in the intermediate group is interesting and worth discussing due to the differences and similarities they show to the other two groups. A future study that specifically focuses on the intermediate group will contribute to understanding the developmental trajectory of L2 English article acquisition by L1 Korean speakers.

5.5 Discussion

The findings of the integration experiment show that L1 speakers process both linguistic and non-linguistic information incrementally, but that advanced L2 speakers, whose overall reaction time does not differ significantly from that of the L1 group, cannot make use of linguistic information when non-linguistic information alone provides accurate predictions for upcoming linguistic material. As for intermediate learners, they showed the tendency to use indefinite articles for singularity marking, which dovetails with the findings of Chapter 4. In this section, I will focus on the patterns displayed by L1 and advanced L2 speakers to discuss how processing efficiency can be an important factor in explaining the discrepancies between L1 and L2 processing.

Both sources of information were integrated incrementally in L1 sentence processing but L2 processing of linguistic information lags behind their processing of non-linguistic information. To understand this claim, two things should be noted: (1) the strengths of the two sources of information as predictive cues and (2) the speed of processing the two types of knowledge. First, for both L1 and advanced L2 speakers, real-world knowledge, limited to the kind in this study,²⁰ elicited a stronger reaction than the definiteness semantics of English articles. Both groups showed a robust main effect of real-world knowledge [$\pm W$] (L1: $\beta = -0.12$, $t\text{-value} = -2.77$, $p = .01$; L2: $\beta = -0.07$, $t\text{-value} = -3.86$, $p = .00$). As can be seen in the visualization of RT change in Figures 5.7, 5.10, and 5.13, real-world knowledge elicits a much more striking

²⁰ One caveat in interpreting this claim is that WK here is limited to common sense shared generally between L1 and L2 adults, the type used in the current dissertation. Real-world knowledge can indicate a wide variety of different types of knowledge and readers should understand that quantum physics is not particularly the type of world knowledge this dissertation aims to look into.

difference than that by grammatical knowledge in both L1 and L2 processing. The amount of RT change between the [+W] and [-W] conditions is much larger than that between the [+G] and [-G] conditions. Along the same vein, when only one of the two sources of information misaligned with the initial expectation, the condition in which parsers' expectation aligned with real-world knowledge [-G+W] induced much faster RTs than that in which it aligned with grammatical knowledge [+G-W]. In brief, to both L1 and L2 speakers, the type of WK used in the current experiment is a much more informative cue than the GK of English articles. This reaction time changes in [\pm G] and [\pm W], therefore, should be interpreted as the extent to which the processing of linguistic and non-linguistic information is efficient or automatized.²¹

The other point is the speed at which the two sources of information are processed or the timeline of when the two groups' reaction to [\pm G] and [\pm W] information appears. L1 speakers show a significant RT difference to the [\pm G] conditions in the [+W] conditions whereas advanced L2 speakers show a significant RT difference induced by [\pm G] in the [-W] conditions. One should pay attention to the fact that [-W] conditions induced slower reaction time than [+W] conditions in both groups. This means that, in L1, linguistic information processing occurs just as quickly and automatically as the processing of [+W] non-linguistic information; however, in L2, the processing of linguistic information is delayed until L2 speakers are surprised by [-W] non-linguistic information and react to it. This, I argue, means that the L2 processing of linguistic information is not as fast or efficient as L1 processing of the same source of

²¹ Here, automatization is defined in general term that most research on automaticity agrees on: The execution of a process does not require any attentional resources and is insuppressible (Kahneman, 1973). That is, looking at the pairing of a doctor and a stethoscope will make it very difficult or impossible not to think about their relationship. The significantly longer reaction time observed in the [-WK] conditions is the evidence that parsers needed the extra time and cognitive efforts to revise their initial expectation.

information. The delayed processing was also observed in Chapter 3 of the current experiment but it means more than just slower processing when another source of information should be processed at the same time.

L2 speakers' insensitivity to linguistic information in the presence of real-world knowledge translates into the possibility that, when non-linguistic information alone can help L2 speakers make sense of a given sentence (i.e., world knowledge alone will allow them to identify a referent yet to be mentioned), they will lose the opportunities to test their linguistic hypotheses. The results clearly show that non-linguistic information, as long as it conforms to one's real-world knowledge, is processed faster than linguistic information in L2 processing. A question that naturally arises at this point is whether L2 speakers will keep processing linguistic information once they reach a good enough resolution of a referent or a good enough interpretation of a sentence based on faster-processed non-linguistic information. The current experimental design does not allow one to make a conclusion on that question. With this empirical question left for future research, one can take a small leap of logic and hypothesize that L2 speakers' reliance on non-linguistic information will lead to missed opportunities of language hypothesis testing.

L2 Interlanguage grammar develops over time (Han, 2004; Selinker, 1972) and, at various points of its development, it is not the same as target grammar. Processing language input (via listening or reading) provides them with opportunities to test if their Interlanguage grammar is the same as target grammar. However, if non-linguistic information gives them cues good enough to interpret a sentence, L2 speakers will not be motivated to continue thinking about grammatical subtleties. This will lead to losing Interlanguage test opportunities, which could be an important reason why the ultimate attainment of L2 acquisition is limited and

oftentimes, many L2 speakers whose length of residence in the L2 environment is as long as 10 years or 59 years (Han, 2004, p. 213) maintain fossilized L2 grammar.

Chapter 6. General Discussion

6.1. Summary

This dissertation investigated how two different sources of information are integrated in L2 sentence processing. The first two experiments tested whether L1 and L2 speakers have the same grammatical knowledge of English articles with respect to definiteness, which is operationalized here as unique identifiability. The third experiment examined whether the two groups share the same type of real-world knowledge (that associates people of certain occupations with certain objects). The final experiment investigated the integration of the two sources of information. The results are summarized below.

Experiment 1. A self-paced reading task probed L1 and L2 speakers' sensitivity to the mapping of unique and non-unique referents to definite and indefinite articles, respectively. L1 speakers showed a reading time increase when the (non-)uniqueness of referents was not aligned with the (in)definiteness of the corresponding referring NPs. In L1 participants, the effects were observed in the spill-over region or one region after the critical NPs were shown. However, advanced L2 speakers displayed the same pattern of behavior one region later than their L1 counterparts. The results suggest that advanced L2 speakers know that the (in)definiteness of English articles indicates the unique identifiability of a referent, although they are slower in applying such knowledge in processing L2 input.

Experiment 2. A referent prediction task inversed the order of information presentation. Unlike Experiment 1, which provided information on the (non-)uniqueness of a referent first and recorded participants' reaction to either matching or mismatching articles, Experiment 2 tested

whether definite or indefinite articles would lead to the prediction of unique or non-unique referents. Incomplete sentences whose final word was a definite or indefinite article were presented along with visual stimuli that had one unique and three non-unique referents. The experiment was designed to see whether either article influences participants' prediction of unique or non-unique referents. Both L1 and advanced L2 speakers predicted unique referents at the cue of a definite article significantly more than at the cue of an indefinite article. Intermediate learners showed the opposite tendency of predicting a unique referent at the cue of an indefinite article numerically more than at a definite article. Although statistically significant effects were not observed in the intermediate group perhaps due to the small sample size, a significant interaction effect of article and speaker group indicated that intermediate learners predict a unique referent significantly more at the cue of an indefinite article than L1 and advanced L2 speakers do at the same cue.

The two takeaway messages from Experiment 2 is that (1) advanced L2 speakers use definite and indefinite articles in a native-like manner to predict unique and non-unique referents and (2) intermediate L2 speakers show the tendency of using indefinite articles for singularity marking instead of as markers of definiteness. There appears to be a native-language effect at work here. Korean, the native language of the L2 participants in this dissertation has no articles and is forced to translate *a/an* as *hana*, the Korean equivalent of *one* in English. A larger sample size in the intermediate group may help confirm this suggestion.

Experiment 3a. A web-based norming survey confirmed that L1 and L2 speakers shared real-world knowledge when it comes to certain associations between people and objects. Some examples of these associations include the relationship between a doctor and a stethoscope, or that between a scientist and a microscope. Both L1 and L2 speakers agreed that doctors are more

strongly associated with stethoscopes than with laptops and that scientists are more closely related to microscopes than to sweaters. Proficiency levels in L2 did not have any effect on the type of world knowledge.

Experiment 3b. As Experiments 1 and 2 confirmed that L1 and advanced L2 speakers shared the same linguistic information with regard to English articles and Experiment 3a demonstrated that non-linguistic information, operationalized as real-world knowledge, was the same for both L1 and L2 groups, it was established that both L1 and L2 speakers have access to the same linguistic and non-linguistic information. With this confirmation, the last experiment looked into the integration of linguistic and non-linguistic information via a referent identification task. The results showed that L1 speakers incrementally integrate both linguistic and non-linguistic information. In particular, they confirm or revise their earlier predictions as they integrate information that becomes available later.

Advanced L2 speakers, who displayed the same capacity for using linguistic and non-linguistic information independently, did not make active use of linguistic information when the sentence corresponded to what they had initially predicted based on their real-world knowledge. When the sentence referred to an object their real-world knowledge hadn't predicted, however, advanced L2 speakers employed linguistic information to revise their initial expectations. This was due to the timing difference of processing linguistic and non-linguistic information. In general, non-linguistic information that matched their real-world knowledge was processed faster than information that is misaligned with their real-world knowledge, and processing linguistic information was as slow as processing non-linguistic information that did not align with real-world knowledge.

The results dovetail with claims that sentence processing is guided not only by linguistic constraints but also by situational, visual, and contextual constraints in L1 processing (Altmann & Kamide, 1999; 2007; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). What was different in L2 sentence processing is that one source of information was processed faster than the other. That is, processing of real-world knowledge was not different in L1 and L2 speakers but L2 processing of the (in)definiteness semantics of English articles was not as automatized as L1 processing of the same information. This difference in the extent to which processing of each information source is automatized was discussed as the potential cause of the fossilization of Interlanguage L2 grammar.

6.2. Theoretical Implications

The recent trend in psycholinguistics literature called for research that investigates the role of processing in understanding language development (Clahsen & Felser, 2006; O'Grady, 2015; Phillips & Ehrenhofer, 2015; Sorace, 2011). Although such research interests were focused on grammatical contingencies were syntax and other linguistic domains (or general cognitive domains) are involved, this dissertation shows that the integration of non-syntactic linguistic information and real-world knowledge provides a good test case for us to understand what sets L2 sentence processing apart from L1 sentence processing. Also, citing O'Grady's work on processing determinism (2015) in relation to the current study needs some clarification in that his focus is not on the variability but on the uniformity of language development.

O'Grady's (2015) processing determinism focuses on how learner-internal and -external processing pressures shapes the development of both first and second language. Processing

determinism's main tenet is that grammar is an epi-phenomenon that reflects parsers' efforts to minimize processing cost. In his Uniformity Thesis, O'Grady argues that the development of language can be highly predictable and uniform when it comes to grammatical contingencies where processing cost is relevant. His discussion of processing cost is certainly not meant to explain variable outcomes of L2 development, not to mention L1.

However, I believe that, based on the findings of the present study, processing determinism can be extended to depict additional sources of variability in L2 behavior (in addition to L1 transfer, which O'Grady (2015) discusses as one of the factors that causes the variability of L2 development). The combination of L2 adults' delayed processing of inflectional morphology and the developed inventory of real-world knowledge leads to the processing cost discrepancy between linguistic and non-linguistic information processing in L2. Learners' resorting to meaning-based, concrete, non-linguistic information routes can be predicted by the thesis that processing efficiency will guide language behavior.

In conclusion, the findings of the current study present that different constraints, which are incrementally integrated in real-time and affect L1 sentence processing, seem to have different timing of integration in L2 processing. This timing difference poses potential challenge in L2 learning in that non-linguistic information processed faster and used for meaning making might demotivate learners to pay attention to, learn, and internalize subtle grammatical tools.

6.3. Pedagogical Implications

The observations in the current dissertation have two meaningful implications with regard to second language pedagogy. One is the developmental pattern of English articles in L1-

Korean L2 speakers of English and the other is the role of real-world knowledge in processing grammar. Although statistical significance was not reached due to its smaller sample size, the results of Experiment 2 (Chapter 4) showed that intermediate L2 learners interpreted indefinite articles as a singularity marker. This was inferred through their predicting unique referents at the cue of an indefinite article. English instruction in Korea, especially in the official curricula of primary and secondary schools, centers around the rote memorization of vocabulary and word-to-word translation. In the process, learners have to learn the meaning of the indefinite article and the first entry for *a/an* in most English-Korean dictionaries is *han(a)* ‘one.’ Although dictionaries do give detailed explanation on how the article is used in various contexts, it is impossible for learners to understand all the functions just by reading the dictionary. They also attempt to translate sentences word by word even when the dictionary interpretation of the indefinite article will not do justice to the actual intention of the writer (or speaker) of the sentence. In circumventing such a negative influence of L1, instructional intervention will help L2 learners learn the accurate grammar.

The current study suggests the possibility that the presence of real-world knowledge could interfere with learners’ paying full attention to grammatical details. This potential interference could lead to the idea that teaching and learning a language in real-life contexts might not be conducive to the development of subtle grammatical features. One cannot deny the effectiveness of communication-based or interaction-based language instructions: The existing literature has been documenting the benefits of language teaching practices such as Content-Based Instruction (Lyster, 2007, Lyster, & Ballinger, 2011; Lyster & Ranta, 1997) and Task-Based Language Teaching (TBLT) (Long, 2016; Van Den Branden, Bygate, and Norris, 2009). In such practices, knowledge external to grammar is used for bootstrapping and enhancing

language development or for making tasks easier, as in Robinson's Cognition Hypothesis in TBLT (Robinson, 2015). What can be tested in pedagogical contexts with regard to the current findings is whether the inclusion or exclusion of real-world knowledge in communicative contexts can make a difference in learning outcomes, especially with regard to subtle grammatical features.

As for content-based instruction (CBI), Roy Lyster mentioned at the post-plenary roundtable after his speech at Second Language Research Forum 2016 that, especially in foreign language settings, CBI can be more effective when lessons cover content that students are not familiar with. If students are already familiar with the subject matter and the focus of the lesson is placed on communicating the knowledge, focus on form, which constitutes the basis of language learning during communicative activities, may not occur as much as desired.

Then, how can one reduce the involvement of real-world knowledge in classrooms where TBLT is the main method of instruction? The essence of TBLT is to construct syllabi and lesson plans based on the results of learners' needs analysis (Long, 2016). That is, it focuses on real-life tasks that learners are most likely to perform in their L2. It may seem counterintuitive to exclude real-world knowledge from tasks in TBLT. However, TBLT employs a great variety of pedagogic tasks in which teachers can modulate the amount of real-world knowledge involved. For example, a problem-solving task can scaffold from familiar topics to unfamiliar ones. If an initial task is to come up with solutions for environmental issues in the city learners are living in, the next task could be about issues in a city that they are not familiar with and the environmental issues could be different from the earlier ones, issues they are not very familiar with.

Another way to maintain language learning activities communicative and keep learners from resorting to their real-world knowledge is to use stories or cultural knowledge they are not

familiar with, e.g., L1-Korean L2 learners of English learning about the customs of American holidays or L1-English L2 learners of Korean learning about the history of the Joseon dynasty (a Korean kingdom that lasted between the 14th and the 19th centuries). It is true that topics that are too unfamiliar might end up demotivating learners but finding the balance between interesting topics to promote communication and novel topics to facilitate grammatical development will be the classroom practitioners' task.

Lastly, I believe the results of the current dissertation provided psycholinguistics grounds for the pedagogical benefits of processing instruction (PI) by Van Patten and his colleagues (Van Patten, 2002; Van Patten, Collopy, Price, Borst, and Qualin, 2013). The basic principles of information processing that Van Patten and colleagues present are that elements of language that contribute to meaning will always be prioritized over those that do not or do so less, and form-related processing will be possible only when meaning-related information can be processed at no cost. The findings of the current dissertation dovetails with the claims of PI: advanced L2 speakers in this study processed meaning-based information (i.e. world knowledge) faster and eventually ended up using information that contributes less to the resolution of a referent (i.e., English articles) only when more meaningful information was unavailable.

6.4. Future research

This dissertation attempted to tease apart linguistic and non-linguistic information and probe their respective and interactional effects on sentence processing. The results are very exciting as well in that learners' language behavior varied in reaction to real-world knowledge. Despite its novel methods and intriguing results, further research is needed to generalize the

findings onto other grammatical structures and other learner groups. Here, I suggest ways in which the experimental design of the current experiment can be extended and/or modified to either generalize the current findings or ask different but relevant research questions: (1) Employing different combinations of information sources, (2) using different online measurements, (3) collecting data from different populations, and (4) applying the insight to pedagogical research.

First of all, the conflict between different sources of information need not be between linguistic and non-linguistic information. The takeaway message of this dissertation is that cues will be prioritized in terms of its strength or effectiveness. Learners will rely on easier-to-process, thus, stronger, cues more than on harder-to-process, thus, weaker, information. Then, the strength of cues will have to be compared in various combinations. For example, the interface hypothesis argues that an external interface can create greater difficulty in L2 learning. However, which information source is easier or harder to process might not depend on whether the sources are within or outside formal grammar. Therefore, testing various combinations of information sources will solidify the argument of the current dissertation that pressure to minimize processing cost can explain second language behavior.

Secondly, eye-tracking experiments can provide time-locked analysis of how each phase of information will influence learners to generate and revise their expectations. Since English articles are one of the most non-salient information phonologically, morphologically, and semantically, the misalignment of articles and visual stimuli might lose its effects towards the end of the experiment as participants get used to the misalignment towards the end of the experiment. This caution was explained in Chapter 4 using Ito, Jincho, Minai, Yamane, and Mazuka (2012) as an example, which reported a main effect of trial blocks. With the report of

the current experiment, it became clear that English articles, despite their subtlety, have a clear enough influence on sentence processing. By observing participants' eye-movements from referent to referent, a clearer picture of initial predictions and revisions (or lack thereof) will shed light on the integration of the two different sources of information.

Next, a similar study on L1 and L2 children will shed light on the validity of the claim that real-world knowledge is the potential cause of the variable outcomes of adult L2 development. It can be assumed that, earlier on the development L1 grammar, children might not have the kind of real-world knowledge that is available to adults. Observing children's behavior at the presence of conflicts between two sources of information will provide a test case to examine the hypothesis that L1 grammar learning precedes almost all other kinds of general learning and that adult learners' efficiency in processing certain types of information backfires on the development of L2 grammar.

Finally, comparing different instructional methods can be another way to test the claim that the presence of real-world knowledge might interfere with mature L2 speakers' grammatical development. Suggestions were made on how different instructional approaches can benefit from modulating the involvement of real-world knowledge (or knowledge familiar and known to learners). Planned research on the effects of real-world knowledge in pedagogical practices such as CBI and TBLT will provide empirical evidence to the claims made in this dissertation.

Appendix 1a: C-test for proficiency

Word Completion Exercise (to be completed within 15 minutes)

Directions: The two texts below contain gaps where parts of some words have been left out (no whole words are missing, though). In the blanks provided, please complete the words so that the sentences and texts make sense. Note that in each blank, you should only complete one word; do not add extra words.

Text 1:

We all live with other people's expectations of us. These are a refle_____ of
th_____ trying to under_____ us; th_____ are predic_____ of
wh_____ they th_____ we will think, d_____ and feel.
Gene_____ we acc_____ the sta_____ quo, but these
expec_____ can be ha_____ to han_____ when they co_____
from our fami_____ and can be diff_____ to ign_____, especially
wh_____ they come from our par_____.

Text 2:

The decision to remove soft drinks from elementary and junior high school vending machines is
a step in the right direction to helping children make better choices when it comes to what they
eat and drink. Childhood obe_____ has bec_____ a ser_____ problem
in th_____ country a_____ children cons_____ more sugar-based
fo_____ and sp_____ less ti_____ getting the nece_____
exercise. Many par_____ have quest_____ schools' deci_____ to
al_____ vending machines which disp_____ candy and so_____
drinks. Many schools, tho_____, have co_____ to re_____ on the
mo_____ these machines generate through agreements with the companies which
makes soft drinks and junk food.

Appendix 1b: Experiment 1– Critical items

‘_’ indicates region boundaries. ‘U’ indicates conditions in which a single unique referent is expected and ‘N’ indicates conditions in which multiple non-unique referents are expected.

Critical NPs were all placed in Region 6.

1. U. In_ the_ kitchen,_ Jason_ checked out_ *the/a stove*_ very carefully_
to make sure_ it was_ safe.
N. In_ the_ appliance store,_ Jason_ checked out_ *the/a stove*_ very carefully_
to make sure_ it was_ safe.
2. U. The student_ walked_ behind the desk_ and sat down_ in_ *the/a chair*_ and
sighed_ rather sadly.
N. The student_ walked_ into the classroom_ and sat down_ in_ *the/a chair*_ and
sighed_ rather sadly.
3. U. At_ his mom's_ birthday party,_ Chris_ admired_ *the/a cake*_ and told_ people_
that he_ liked it.
N. In_ the fancy_ bakery,_ Chris_ admired_ *the/a cake*_ and told_ people_
that he_ liked it.
4. U. In_ her_ living room,_ Susan_ turned on_ *the/a flat-screen TV*_ but was not
pleased_ with the picture quality.
N. In_ the_ electronics store,_ Susan_ turned on_ *the/a flat-screen TV*_ but was not
pleased_ with the picture quality.
5. U. In her_ basement,_ Becky heard_ a bad noise_ coming from_
*the/a washing machine*_ and started_ worrying that_ the machine was_ broken.

- N. In the _ laundromat, _ Becky heard _ a bad noise _ coming from _
the/a washing machine _ and started _ worrying that _ the machine was _ broken.
6. U. As soon as _ he walked _ into his biology class, _ Zachary _ asked _
the/a professor _ how long _ her conference trip _ would _ be.
- N. As soon as _ he walked _ into the faculty lounge, _ Zachary _ asked _
the/a professor _ how long _ her conference trip _ would _ be.
7. U. Picking up _ her hamburger, _ Patty noticed _ mold _ on _ *the/a bun* _ so
she _ had to _ throw it _ away.
- N. Making hamburgers _ at McDonalds, _ Patty noticed _ mold _ on _ *the/a bun* _ so
she _ had to _ throw it _ away.
8. U. Looking at _ the parked car, _ Frank thought _ he _ recognized _ *the/a driver* _ but
realized _ he was _ wrong.
- N. Looking at _ the line of cabs, _ Frank thought _ he _ recognized _ *the/a driver* _ but
realized _ he was _ wrong.
9. U. Putting _ the single flower _ into the vase, _ Abby had to _ trim _
the/a stem _ because _ the vase was _ short.
- N. Putting _ the bunch of flowers _ into the vase, _ Abby had to _ trim _
the/a stem _ because _ the vase was _ short.
10. U. In _ her _ bathroom _ at home, _ Sally cleaned _
the/a toilet _ squeaky _ clean _ with _ a bristle brush.
- N. In _ the _ Campus Center _ women's room, _ Sally cleaned _
the/a toilet _ squeaky _ clean _ with _ a bristle brush.

11. U. When he_ woke up_ in the morning,_ John_ picked up_ *the/a phone*_ and called_ his_ girlfriend.
- N. When he_ went_ to the AT&T shop,_ John_ picked up_ *the/a phone*_ and checked_ its_ price.
12. U. In her_ bedroom,_ she_ threw herself_ on_ *the/a bed*_ and_ fell asleep_ before she_ knew it.
- N. In the_ furniture store,_ she_ threw herself_ on_ *the/a bed*_ and_ fell asleep_ before she_ knew it.
13. U. In John's_ math class_ today,_ someone threw_ a pen at_ *the/a teacher*_ and got suspended_ for two weeks.
- N. At John's_ high school_ today,_ someone threw_ a pen at_ *the/a teacher*_ and got suspended_ for two weeks.
14. U. Typing homework_ on his home computer,_ George_ spilled_ coffee on_ *the/a keyboard*_ and it_ stopped_ working.
- N. Looking around_ at the Apple store,_ George_ spilled_ coffee on_ *the/a keyboard*_ and it_ stopped_ working.
15. U. To reschedule_ his appointment_ at the clinic,_ Dan_ called_ *the/a therapist*_ to see_ what times_ were available.
- N. To schedule_ a massage_ for his broken leg,_ Dan_ called_ *the/a therapist*_ to see_ what times_ were available.
16. U. Right after the couple_ walked into_ the empty waiting room,_ the woman started to_ scream at_ *the/a man*_ very loudly_ and scared_ people_ in the hallway.

- N. Right after a group of people_ walked into_ the empty waiting room,_ a woman started to_ scream at_ ***the/a man***_ very loudly_ and scared_ people_ in the hallway.
17. U. As she finished_ folding the shirt,_ Jane realized that_ she hadn't_ ironed_ ***the/a collar***_ so she took out_ the ironing board_ again.
- N. As she finished_ folding her shirts,_ Jane realized that_ she hadn't_ ironed_ ***the/a collar***_ so she took out_ the ironing board_ again.
18. U. After_ pouring coffee_ into her tumbler,_ Elaine_ put_ ***the/a lid***_ on it_ and_ handed it_ to her husband.
- N. After_ pouring coffee_ into two tumblers,_ Elaine_ put_ ***the/a lid***_ on one_ and_ handed it_ to her husband.
19. U. Passing_ by_ the local deli,_ Molly_ greeted_ ***the/a shop owner***_ and asked_ about_ his wife's_ health.
- N. Walking_ around_ the marketplace,_ Molly_ greeted_ ***the/a shop owner***_ and asked_ about_ his wife's_ health.
20. U. Opening_ the new laptop computer box,_ Katrina_ accidentally_ damaged_ ***the/a monitor***_ so she_ had to ask_ her friend_ for help.
- N. Handling_ the lab equipment,_ Katrina_ accidentally_ damaged_ ***the/a monitor***_ so she_ had to ask_ her friend_ for help.

Appendix 1c: Experiment 1– Filler items

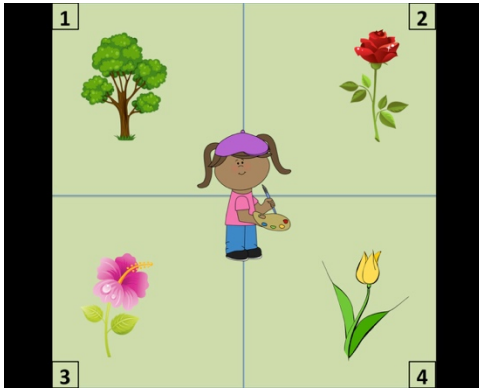
1. At_ the_ lunch table_ the children_ fought over_ the last cupcake_ but a bully_
showed up_ and took it_ away.
2. At_ the_ Birch Run Mall_ Miranda bought_ shirts_ and skirts_ for herself_
and her sister.
3. At_ the_ hotel front desk_ Sally asked for_ information_ before she_ went out_
sightseeing.
4. On_ Main Street_ the tourists_ saw_ Justin Bieber_ shooting a_ commercial_ for Macy's.
5. On_ the_ bottom_ of the page_ Jimmy wrote_ his name_ and number_ to join the
volunteer group.
6. On_ top_ of the_ Christmas tree_ was_ an_ angel_ that shone_ like_ a star.
7. On_ every_ table_ in the law library_ is a sign_ that says_ "Reserved_ for law students."
8. In_ the_ master bedroom_ Jack and_ his wife_ talked_ about politics_
for hours and hours.
9. After_ finishing_ lunch,_ Melody read_ a novel_ and took a nap_ until James_
came home.
10. Because_ Anna_ was_ on_ a_ diet,_ she only had_ five cookies_ instead of_
her usual ten.
11. While_ shopping in_ the mall,_ Claire saw_ the student_ who had_ gone home_ early_
because of a stomach flu.
12. Although_ Mike_ left_ Hawaii,_ he missed_ people_ he had met there_ and wanted to_
go back.

13. Before_ the building_ was renovated,_ it looked like_ it was haunted_ and no one_ was_ living there.
14. While_ the plumber_ was_ getting a tool_ from his truck,_ the toilet exploded_ and flooded_ the bathroom.
15. It_ was_ surprising_ that_ Ben_ had never_ been_ to that restaurant_ since he said_ he liked Thai food.
16. Yesterday,_ no one knew_ why Julia_ didn't show up_ at the meeting,_ but this morning_ she_ called in sick.
17. Jonathan was_ watching_ a reality show_ on TV_ when his friend_ called_ to invite him over_ for dinner.
18. The professor_ scheduled_ a make-up exam_ because_ a third of her students_ failed_ the_ original exam.
19. The box_ on_ the_ table_ was delivered_ this morning_ about an hour_ after Jennifer left_ for the library._
20. The_ policeman arrived_ too late_ so_ the_ thief had_ had plenty of time_ to flee the scene.
21. Several_ women_ in fancy clothes_ came to the store_ and purchased_ many purses and_ pairs of_ shoes.
22. The sun_ was setting_ and_ the traffic_ became heavy_ as Hannah was_ sitting_ in her car_ listening to_ the radio.
23. The girl_ with_ red hair_ yelled_ at the man_ to leave_ the womens lounge;_ soon,_ campus security_ arrived.

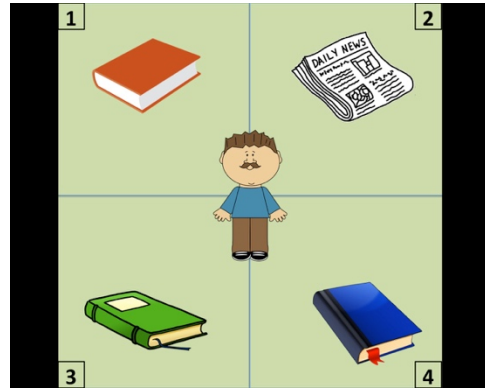
24. It was _ a _ pleasant _ surprise _ to hear from _ an old friend _ of mine _ who had been _ out of touch _ for a while.
25. Next week _ all students _ enrolled in _ Physics 101 _ have to attend _ lab sessions with _ a designated _ teaching assistant.
26. On Wednesday _ there will be _ a small gathering _ of our friends _ at _ a _ nearby _ restaurant.
27. Last weekend _ Jared and _ his friends _ went camping _ at the park _ located at _ the northern end _ of Kaneohe.
28. In 1982 _ Michael Jackson's _ album Thriller _ was released, _ and became _ the best-selling _ album _ ever.
29. Yesterday _ Ala Moana _ Beach Park _ was full _ of people _ trying to _ secure a spot _ for the lantern festival.
30. Tomorrow night _ a farewell party _ for _ John _ and _ Alex _ will be held _ at Melanie's place.
31. Right _ after _ breakfast _ a postman _ knocked _ on our door _ and delivered _ a package _ from _ my parents.
32. On _ a _ mid-summer _ night _ the boy _ kissed the girl _ under the moon _ and sang her _ a love song.
33. To remove _ a stain _ on her new shirt _ Bonnie used _ chlorine bleach, _ but it wasn't _ very _ effective.
34. To improve _ his Spanish _ Jake _ has decided to _ travel to Spain _ this summer, _ and so he is _ saving _ his money.

35. To increase_ his chance_ of promotion,_ Kirk played_ golf_ with his boss_ and everyone_ hated him_ for it.
36. To visit_ her grandmother_ in the hospital,_ Amanda took_ a week off work_ and drove_ to_ California.
37. To make_ Sarah_ jealous,_ an immature boy_ tried_ to flirt with every girl_ in sight_ but to no avail.
38. To make up for_ the missed class,_ the teacher_ had to_ give an extra lecture_ before_ the final_ exam.
39. In order to_ go hiking,_ Laura_ bought_ a new pair of boots,_ but the trip was_ cancelled_ due to_ bad weather.
40. The author_ of the book_ used_ real life stories_ from people he met_ during his travels_ in_ Asia.

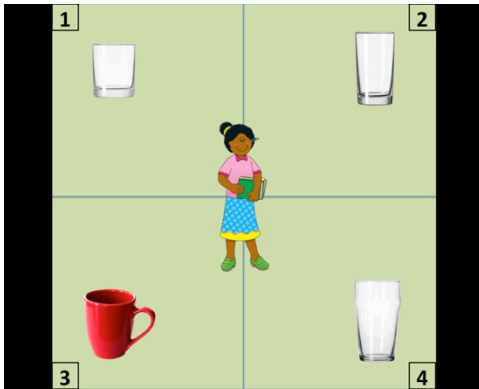
Appendix 2a: Experiment 2–Critical items



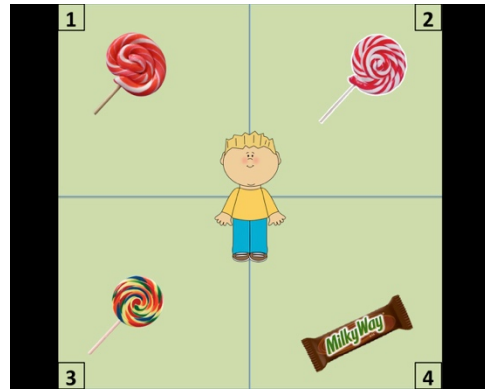
1. The painter will want to draw the/a...



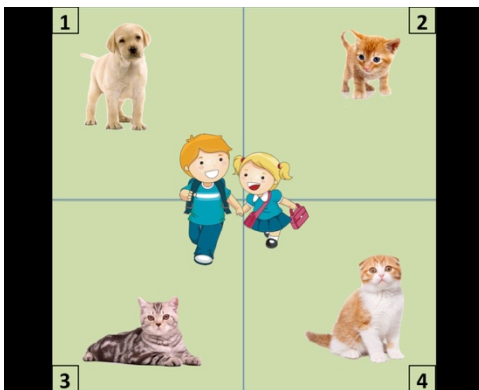
2. The man will read the/a...



3. The woman will buy the/a...



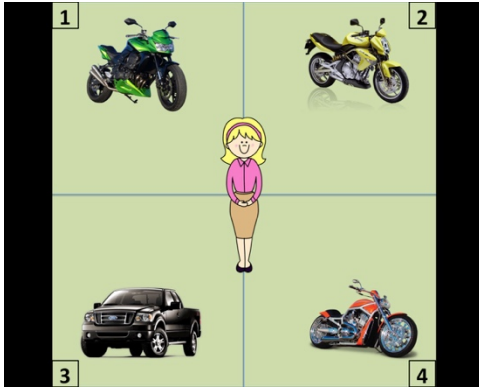
4. The child will eat the/a...



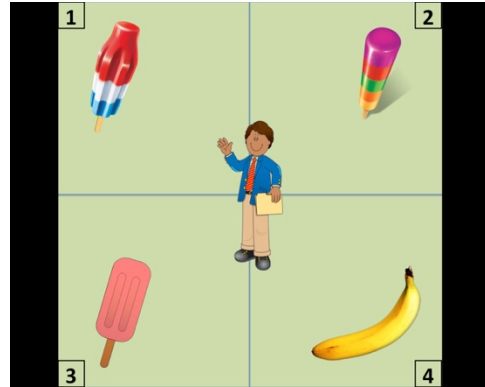
5. The children will buy the/a...



6. The lady will choose the/a...



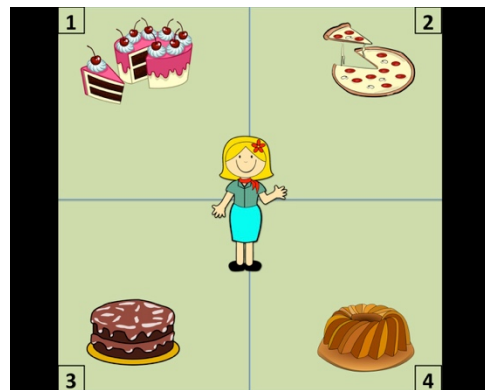
7. The woman will test-drive the/a...



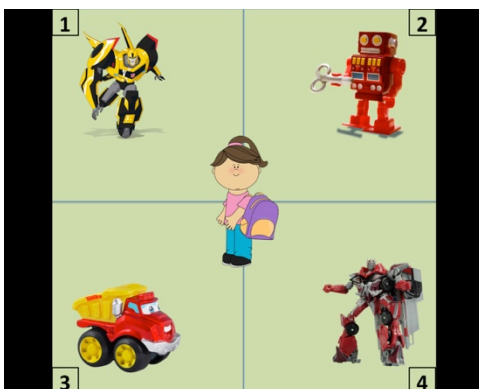
8. The man will taste the/a...



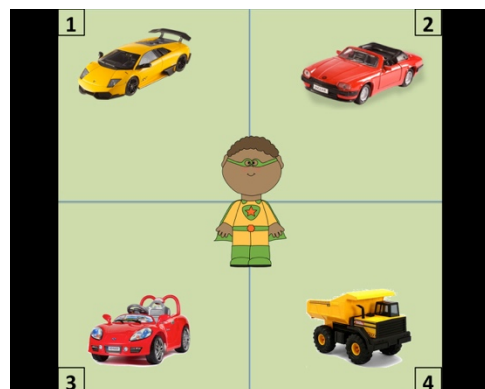
9. The soldier will photograph the/a...



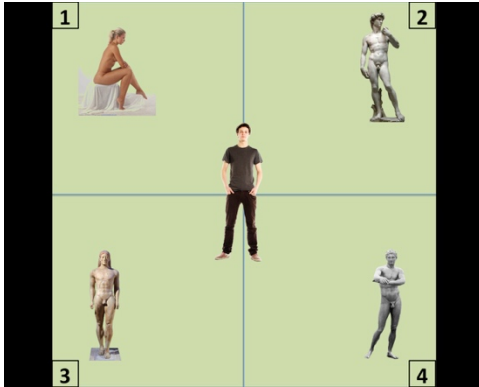
10. The customer will order the/a...



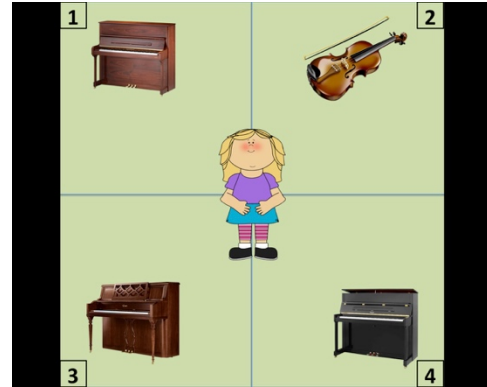
11. The girl will select the/a...



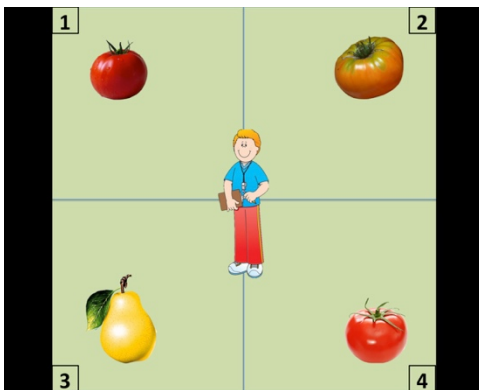
12. The boy will purchase the/a...



13. The painter will paint the/a...



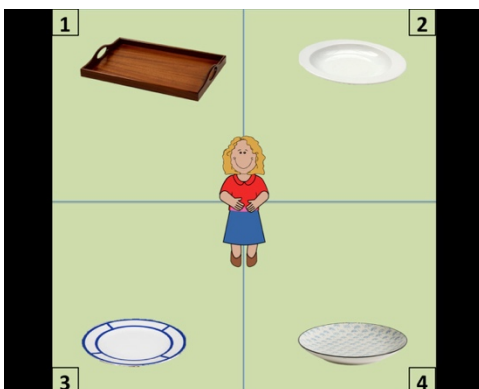
14. The kid will try the/a...



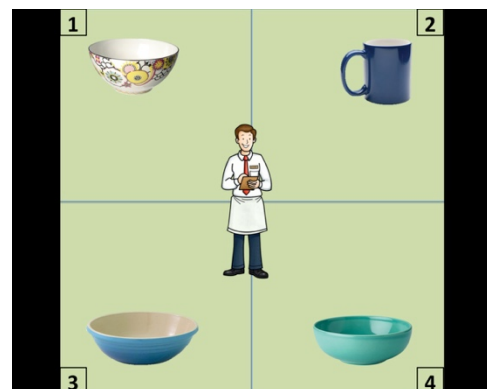
15. The man will throw the/a...



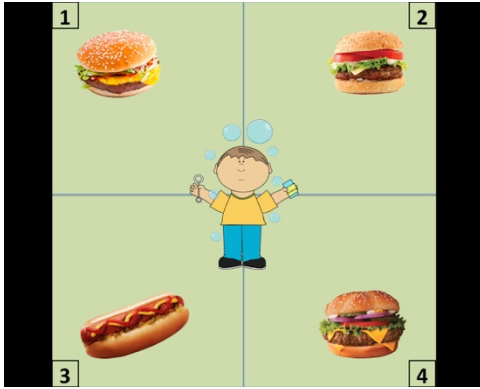
16. The woman will toss the/a...



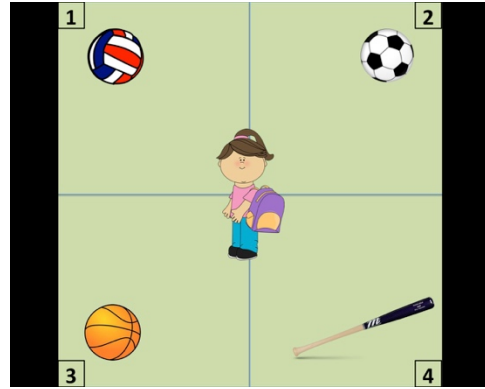
17. The lady will carry the/a...



18. The guy will wash the/a...



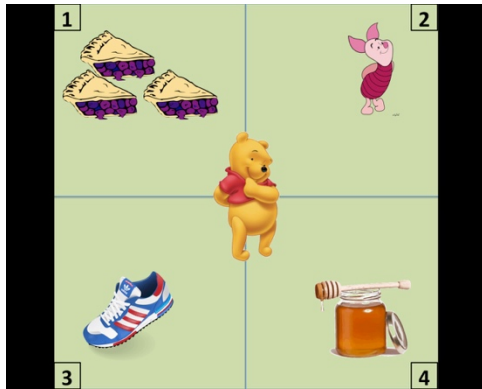
19. The boy will take the/a...



20. The girl will pick the/a...

Appendix 2b: Experiment 2–Filler samples

One-referent filler item



The Pooh bear will get a piece of...

Four-referent filler item



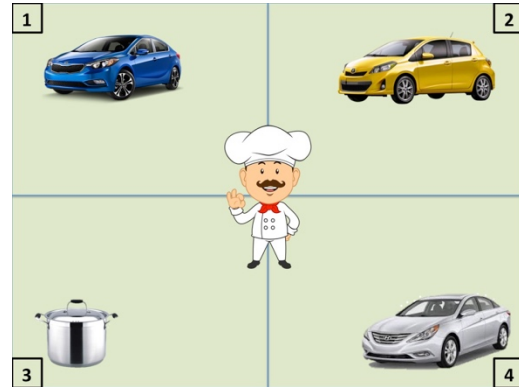
The tiger will drink some...

Appendix 3a: Experiment 3–Critical items



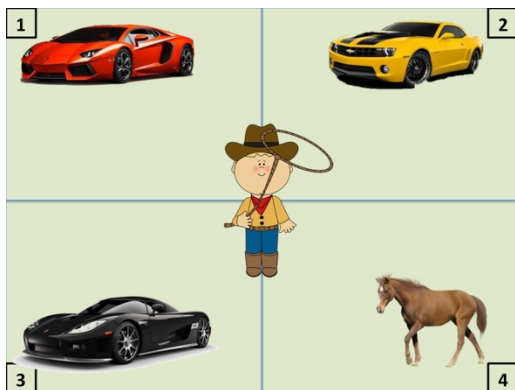
1. The baby wants to get the/a pacifier

Barbie doll



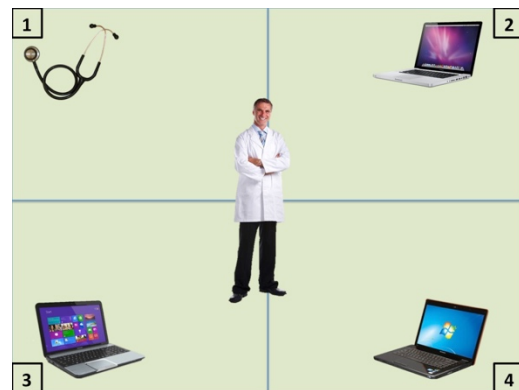
2. The cook will probably buy the/a pot

car



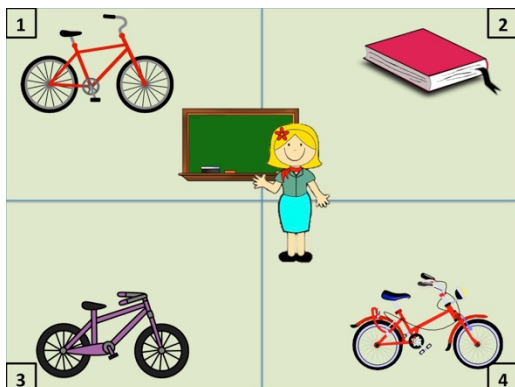
3. The boy will want to ride the/a horse

car



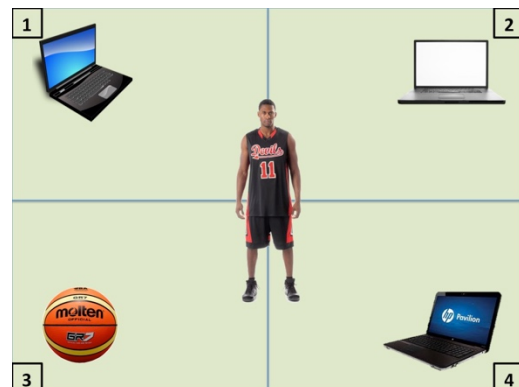
4. The man will want to use the/a stethoscope

laptop



5. She will probably want the/a book

bike

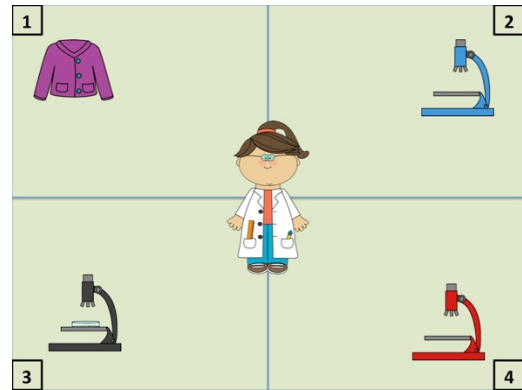


6. The man will probably want the/a basketball

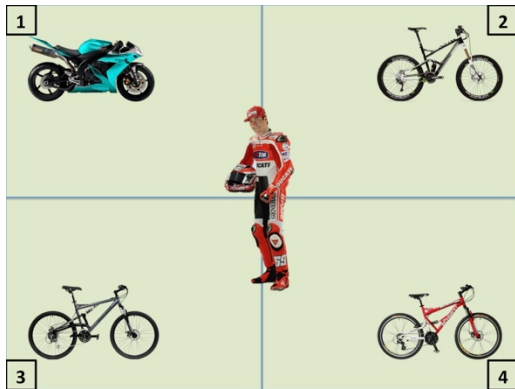
computer



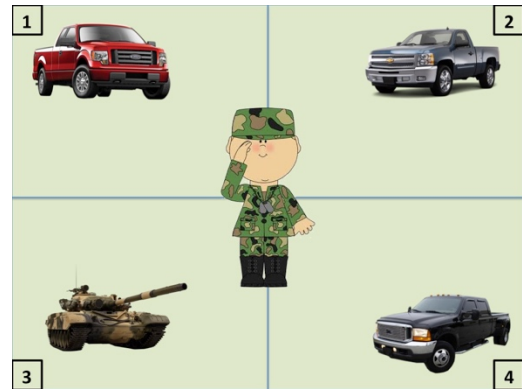
7. The woman will want to purchase the/a suitcase
mug



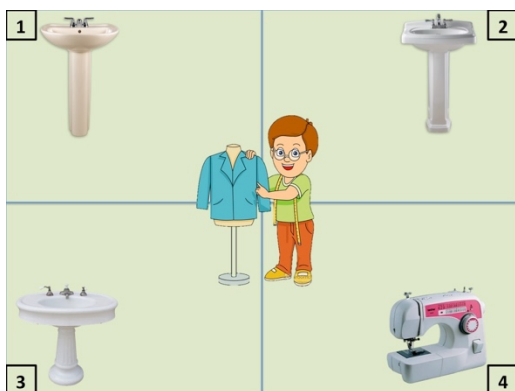
8. She will want to get the/a microscope
sweater



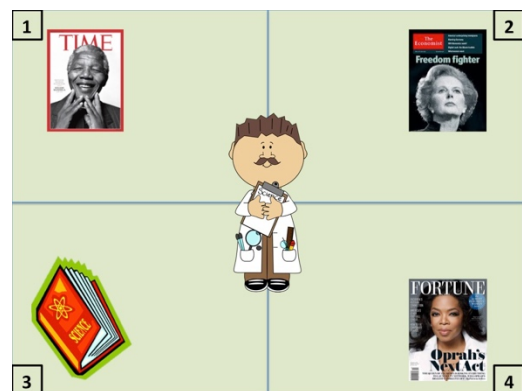
9. The man will want to ride the/a motorcycle
bicycle



10. He will want to ride the/a tank
truck



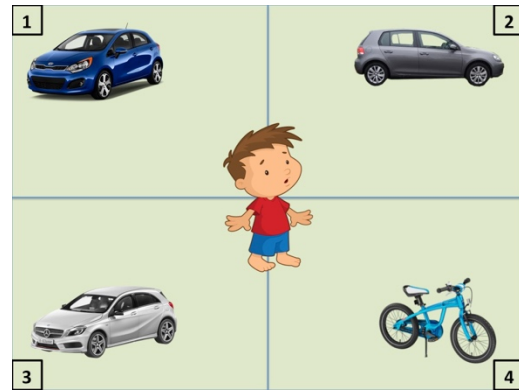
11. The man will want to repair the/a sewing machine
bathroom sink



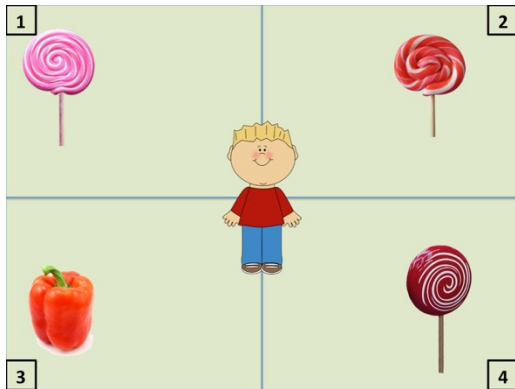
12. He will want to read the/a science book
magazine



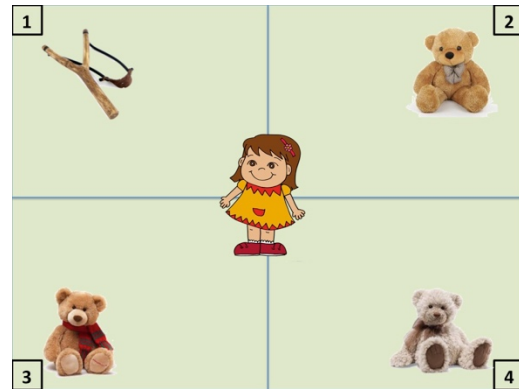
13. She will probably want the/a handbag
briefcase



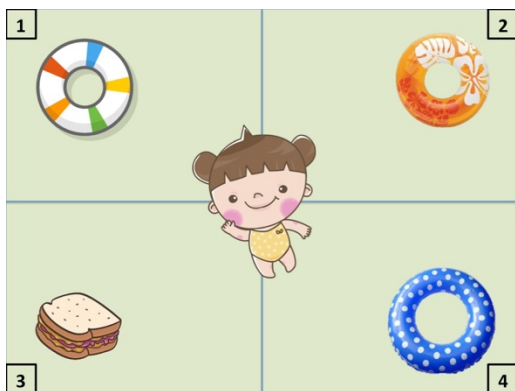
14. The boy will want to ride the/a bike
car



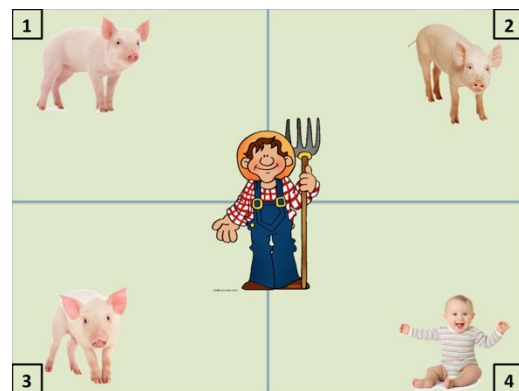
15. The child wants to taste the/a lollipop
bell pepper



16. The girl will want to get the/a teddy bear
slingshot



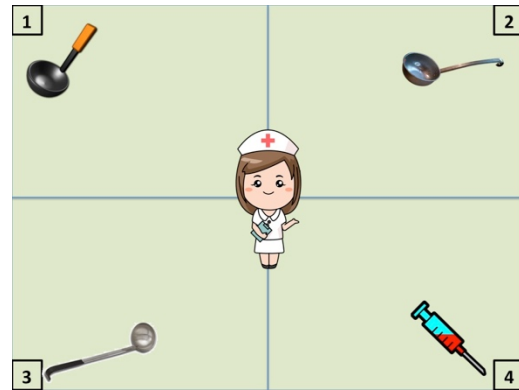
17. The girl will probably want the/a inner tube
sandwich



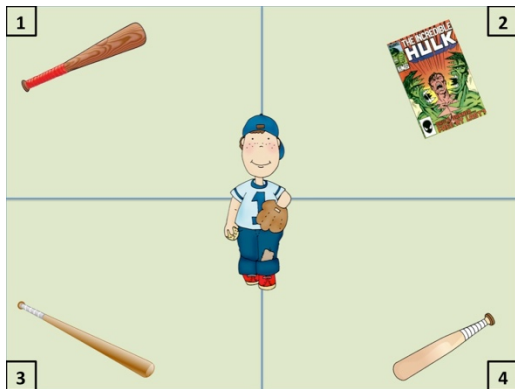
18. The man will probably feed the/a pig
baby



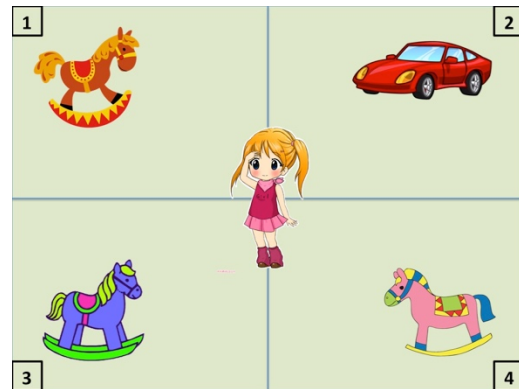
19. The man will want to get the/a hammer
ball



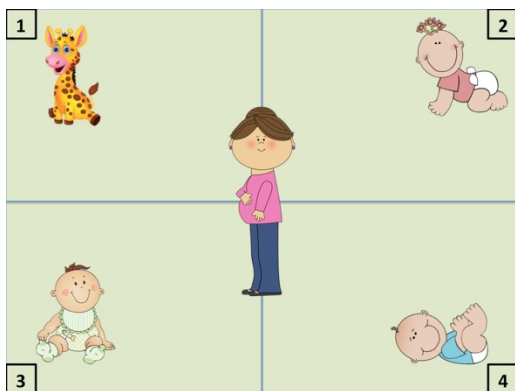
20. She will want to use the/a syringe
ladle



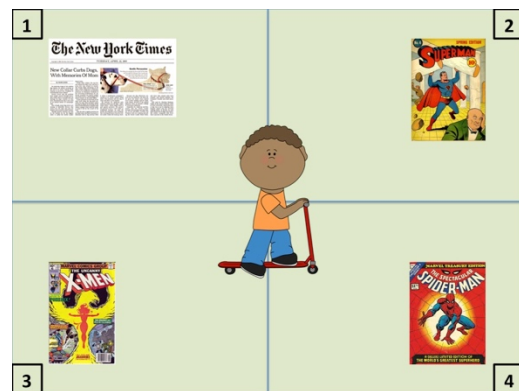
21. The boy will want to get the/a baseball bat
comic book



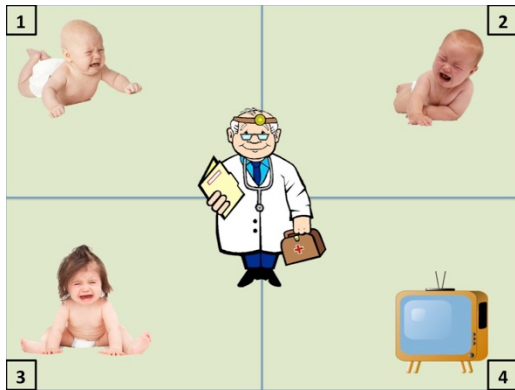
22. The girl will want to ride the/a horse
car



23. She will feed the/a baby
giraffe

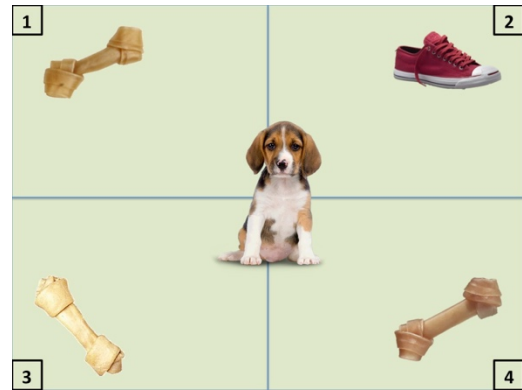


24. He will want to read the/a comic book
newspaper



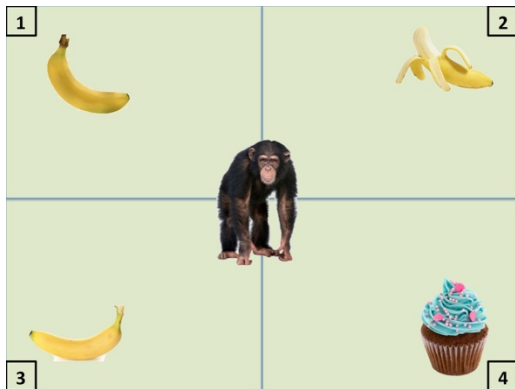
25. The man will examine the/a baby

TV



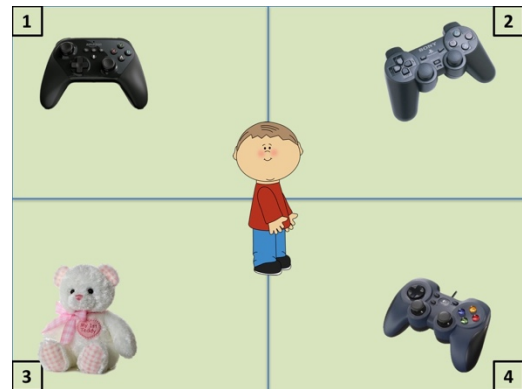
26. The puppy will want to choose on the/a bone

shoe



27. The chimpanzee will want to eat the/a banana

cupcake



28. The child will want to play with the/a Gameboy

teddy bear

Appendix 3b–Experiment 3: Filler samples

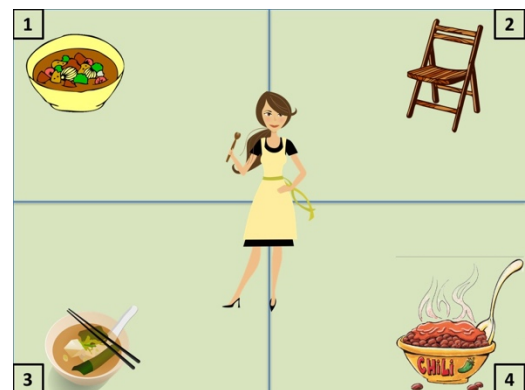
1. One-referent filler item (x 7)

The baby will need to eat baby food.



2. Two-referent filler item (x 21)

The child will put on his shirt.



3. Three-referent filler item

The girl will want to make her dinner.



4. Four-referent filler item

The child will want to have some drink.

References

- Abbott, B. (2006). Definiteness and indefiniteness. In L. R. Horn & G. Ward (Eds.), *The handbook of pragmatics* (pp. 122-150). Malden, MA: Blackwell publishing.
- Ahern, C., & Stevens, J. (2014). How uniqueness guides definite description processing. *UPenn Working Papers in Linguistics*, 20(1), 1–10.
- Ahn, H. (2009). English article use by non-native English teachers in Korea: Task-type variability in interlanguage performance and its pedagogical implications. *SNU Working Papers in English Linguistics and Language*, 8, 43–58.
- Altmann, G. T. (1999). Thematic role assignment in context. *Journal of Memory and Language*, 41(1), 124–145.
- Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition*, 73(3), 247–264.
- Altmann, G. T. M., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, 57(4), 502-518.
- Altmann, G. T., & Kamide, Y. (2009). Discourse-mediation of the mapping between language and the visual world: eye movements and mental representation. *Cognition*, 111(1), 55-71.
- Aronson, D. & Scarborough, H. S. (1976). Performance theories for sentence coding: Some quantitative evidence. *Journal of Experimental Psychology: Human perception and performance*, 2(1), 56–70.

- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278.
- Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. Retrieved from <https://arxiv.org/abs/1506.04967>
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- Bickerton, D. (1981). *Roots of language*. Ann Arbor: Karoma Publishers.
- Blom, E., Polisenska, D., & Weerman, F. (2008). Articles, adjectives and age of onset: the acquisition of Dutch grammatical gender. *Second Language Research*, 24(3), 297–331.
- Boersma, Paul & Weenink, David (2016). Praat: doing phonetics by computer [Computer program]. Version 6.0.15, retrieved 23 March 2016 from <http://www.praat.org/>
- Bohnacker, U. (1997). Determiner Phrases and the Debate on Functional Categories in Early Child Language. *Language Acquisition*, 6(1), 49–90.
- Brown, J. D. (1997). Skewness and kurtosis. *JALT Testing & Evaluation SIG Newsletter*, 1(1), 90–23.
- Burkhardt, P. (2006). Inferential bridging relations reveal distinct neural mechanisms: evidence from event-related brain potentials. *Brain Lang*, 98(2), 159–168.
- Burkhardt, P. (2008a). Two types of definites: evidence for presupposition cost. In A. Grønn (Ed.), *Proceedings of Sinn und Bedeutung 12* (pp. 66–80). Oslo: Department of Literature, Area Studies and European Languages, University of Oslo.

- Burkhardt, P. (2008b). Dependency precedes independence: Online evidence from discourse processing. In A. Benz & P. Knhlein (Eds.), *Constraints in discourse* (pp. 141–158). Amsterdam: John Benjamins Publishing Company.
- Butler, Y. G. (2002). Second language learners' theories on the use of English articles. *Studies in Second Language Acquisition*, 24(3), 451–480.
- Chambers, C. G., Tanenhaus, M. K., Eberhard, K. M., Filip, H., & Carlson, G. N. (2002). Circumscribing Referential Domains during Real-Time Language Comprehension. *Journal of Memory and Language*, 47(1), 30–49.
- Chambers, C. G., Tanenhaus, M. K., & Magnuson, J. S. (2004). Actions and affordances in syntactic ambiguity resolution. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(3), 687-696.
- Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology*, 42(4), 368-407.
- Christianson, K., Williams, C. C., Zacks, R. T., & Ferreira, F. (2006). Younger and Older Adults' "Good-Enough" Interpretations of Garden-Path Sentences. *Discourse Process*, 42(2), 205–238.
- Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, 27(1), 3–42.
- Clifton Jr., C. (2013). Situational context affects definiteness preferences: Accommodation of presuppositions. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 39(2), 487–501.

- Clifton Jr, C., & Frazier, L. (1989). Comprehending sentences with long-distance dependencies. In G. N. Carlson & M. K. Tanenhaus (Eds.), *Linguistic structure in language processing* (pp. 273-317). Dordrecht, the Netherlands: Springer.
- Clifton, C. J., Traxler, M. J., Mohamed, M. T., Williams, R. S., Morris, R. K., & Rayner, K. (2003). The use of thematic role information in parsing: Syntactic processing autonomy revisited. *Journal of Memory and Language*, 49(3), 317–334.
- Cook, A. E., & Myers, J. L. (2004). Processing discourse roles in scripted narratives: The influences of context and world knowledge. *Journal of Memory and Language*, 50(3), 268–288.
- Crawley, M. (2007). *The R Book*. Chichester, England: John Wiley.
- Dekydtspotter, L., Schwartz, B. D., & Sprouse, R. A. (2006). The comparative fallacy in L2 processing research. In M. G. O'Brien, C. Shea, & J. Archibald (Eds.), *The 8th Generative Approaches to Second Language Acquisition* (pp. 33–40). Sommerville, MA: Cascadilla Proceedings Project.
- Demuth, K. (2001a). A prosodic approach to filler syllables. *Journal of Child Language*, 28(1), 246–249.
- Demuth, K. (2001b). Prosodic constraints on morphological development. In J. Weissenborn & B. Hohle (Eds.), *Approaches to bootstrapping: Phonological, lexical, syntactic and neurophysiological aspects of early language acquisition* (Vol. 2, pp. 3–21). Amsterdam/Philadelphia: John Benjamins.
- Ekiert, M. (2004). Acquisition of the English article system by speakers of Polish in ESL and EFL settings *Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics* (Vol. 4, pp. 1–23).

- Ekiert, M. B. (2010). *Articles as expressions of definiteness in L2 English of Slavic speakers*. (Doctoral Dissertation), Teachers College, Columbia University, New York.
- Ferreira, F. (2003). The misinterpretation of non-canonical sentences. *Cognitive Psychology*, 47(2), 164–203.
- Ferreira, C., Bailey, K. G. D., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current directions in psychological science*, 11(1), 11–15.
- Ferreira, F., Engelhardt, P. E., & Jones, M. W. (2009). Good enough language processing: a satisficing approach. In N. Taatgen, H. Rijn, J. Nerbonne, & L. Schomaker (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 413–418). Austin, TX: Cognitive Science Society.
- Ferreira, F., & Patson, N. D. (2007). The 'good enough' approach to language comprehension. *Language and Linguistics Compass*, 1(1-2), 71–83.
- Fodor, J. D. (1998). Learning to parse. *J Psycholinguist Res*, 27(2), 285–319.
- Fodor, J. D. (1998). Parsing to learn. *Journal of Psycholinguist Research*, 27(3), 339–374.
- Garcia Mayo, M. d. P. (2009). Article choice in L2 English by Spanish speakers. In M. d. P. Garcia Mayo & R. Hawkins (Eds.), *Second language acquisition of articles: Empirical findings and theoretical implications* (pp. 13–35). Philadelphia: John Benjamins.
- Gerken, L. (1991). The metrical basis for children's subjectless sentences. *Journal of Memory and Language*, 30(4), 431–451.
- Gerken, L. (1994). Young children's representation of prosodic phonology: Evidence from English-speakers' weak syllable productions. *Journal of Memory and Language*, 33(1), 19–38.
- Gerken, L. (1996a). Prosodic structure in young children's language production. *Language*,

72(4), 683–712.

Gerken, L. (1996b). Prosody's role in language acquisition and adult parsing. *Journal of Psycholinguistic Research*, 25(2), 345–356.

Gerken, L., & Landau, B. (1990). Function morphemes in young children's speech perception and production. *Developmental psychology*, 26(2), 204-216.

Gerken, L., & McIntosh, B. J. (1993). Interplay of function morphemes and prosody in early language. *Developmental psychology*, 29(3), 448-457.

Goad, H., & Buckley, M. (2006). Prosodic structure in child French: evidence for the foot. *Canadian Journal of Linguistics*, 5, 109-142.

Goad, H., & White, L. (2009). Prosodic transfer and the representation of determiners in Turkish-English interlanguage. In N. Snape, Y.-k. I. Leung, & M. Sharwood Smith (Eds.), *Representational Deficits in SLA: Studies in honor of Roger Hawkins* (pp. 1–26). Amsterdam: Johns Benjamins Publishing Company.

Grüter, T., Lew-Williams, C., & Fernald, A. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research*, 28(2), 191-215.

Guasti, M. T., Gavarró, A., de Lange, J., & Caprin, C. (2008). Article Omission Across Child Languages. *Language Acquisition*, 15(2), 89–119.

Gürel, A. (2011). In search for a unified model of L2 acquisition and L1 attrition: A commentary for the Interface Hypothesis. *Linguistic Approaches to Bilingualism*, 1(1), 39–42.

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, 304(5669), 438–441.

Han, Z. (2004). Fossilization: five central issues. *International Journal of Applied Linguistics*, 14(2), 212–242.

- Hawkins, J. A. (1991). On (In)definite Articles: Implicatures and (un)grammaticality prediction. *Journal of Linguistics*, 27(2), 405-442.
- Heim, I. (1988). *The semantics of definite and indefinite noun phrases*. New York: Garland Publishing Inc.
- Hopp, H. (2016). Learning (not) to predict: Grammatical gender processing in second language acquisition. *Second Language Research*, 32(2), 277-307.
- Huebner, T. (1979). Order-of-acquisition vs. dynamic paradigm: a comparison of method in interlanguage. *TESOL Quarterly*, 13(1), 21–28.
- Huebner, T. (1983) *A longitudinal analysis of the acquisition of English*. Ann Arbor, MI: Karoma.
- Ionin, T. (2016). What do L2-English learners know about the semantics of D? [PDF slides prepared for the Dimensions of D Workshop at the University of Rochester]. Obtained from the author.
- Ionin, T., Baek, S., Kim, E., Ko, H., & Wexler, K. (2012). That's not so different from the: Definite and demonstrative descriptions in second language acquisition. *Second Language Research*, 28(1), 69–101.
- Ionin, T., Kim, M.-H., Tyndall, K. (in progress). Targeting explicit vs. implicit knowledge of English articles.
- Ionin, T., Ko, H., & Wexler, K. (2004). Article semantics in L2 acquisition: The role of specificity. *Language Acquisition*, 12(1), 3–69.
- Ionin, T., Ko, H., & Wexler, K. (2007). The role of semantic features in the acquisition of English articles by Russian and Korean speakers. In J. M. Liceras, H. Zobl, & H. Goodluck (Eds.), *The Role of Formal Features in Second Language Acquisition*. London:

Routledge.

- Ionin, T., Zubizarreta, M. L., & Maldonado, S. B. (2008). Sources of linguistic knowledge in the second language acquisition of English articles. *Lingua*, 118(4), 554–576.
- Ionin, T., Zubizarreta, M. L., & Philippov, V. (2009). Acquisition of article semantics by child and adult L2-English learners. *Bilingualism: Language and Cognition*, 12(03), 337–361.
- Ito, K., Jincho, N., Minai, U., Yamane, N., & Mazuka, R. (2012). Intonation facilitates contrast resolution: Evidence from Japanese adults and 6-year olds. *Journal of Memory and Language*, 66(1), 265–284.
- Jaeger, T. F. (2008). Categorical Data Analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of Memory and Language*, 59(4), 434–446.
- Jarvis, S. (2002). Topic continuity in L2 English article use. *Studies in Second Language Acquisition*, 24, 387–418.
- Juffs, A., & Harrington, M. (1995). Parsing effects in second language sentence processing. *Studies in Second Language Acquisition*, 17(4), 483–516.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.
- Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, 49(1), 133–156.
- Kamide, Y., Scheepers, C., & Altmann, G. T. (2003). Integration of syntactic and semantic information in predictive processing: cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research*, 32(1), 37–55.
- Karimi, H., & Ferreira, F. (2016). Good-enough linguistic representations and online cognitive equilibrium in language processing. *The Quarterly Journal of Experimental Psychology*,

69(5), 1013–1040.

- Kim, L. K., & Lakshmanan, U. (2009). The processing role of the article choice parameter: Evidence from L2 learners of English. In M. d. P. Garcia Mayo & R. Hawkins (Eds.), *Second language acquisition of articles: Empirical findings and theoretical implications*. Amsterdam: John Benjamins.
- Ko, H., Ionin, T., & Wexler, K. (2004). Adult L2-learners lack the maximality presupposition, too! In K. U. Deen, J. Nomura, B. Schulz, & B. D. Schwartz (Eds.), *Proceedings of the Inaugural Conference on Generative Approaches to Language Acquisition-North America (GALANA 1)* (pp. 171–182). Honolulu, HI.
- Ko, H., Ionin, T., & Wexler, K. (2010). The role of presuppositionality in the second language acquisition of English articles. *Linguistic Inquiry*, 41(2), 213–254.
- Kupisch, T., Anderssen, M., Bohnacker, U., & Snape, N. (2009). Article acquisition in English, German, Norwegian, and Swedish. In R. P. Leow, H. Campos, & D. Lardiere (Eds.), *Little words: Their history, phonology, syntax, semantics, pragmatics, and acquisition*. Washington, DC: Georgetown University Press.
- Lardiere, D. (1998). Dissociating syntax from morphology in a divergent L2 end-state grammar. *Second language research*, 14(4), 359-375.
- Lew-Williams, C., & Fernald, A. (2010). Real-time processing of gender-marked articles by native and non-native Spanish speakers. *Journal of Memory and Language*, 63(4), 447-464.
- Liu, D., & Gleason, J. L. (2002). Acquisition of the article *the* by nonnative speakers of English: An analysis of four non-generic uses. *Studies in Second Language Acquisition*, 24, 1–26.
- Long, M. H. (2016). In Defense of Tasks and TBLT: Nonissues and Real Issues. *Annual Review*

- of Applied Linguistics*, 36, 5-33.
- Lyons, C. (1999). *Definiteness*. Cambridge: Cambridge University Press.
- Lyster, R. (2007). *Learning and teaching languages through content: A counterbalanced approach*. Amsterdam/Philadelphia: John Benjamins Publishing.
- Lyster, R., & Ballinger, S. (2011). Content-based language teaching: Convergent concerns across divergent contexts. *Language Teaching Research*.
- Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake. *Studies in Second Language Acquisition*, 19(1), 37–66.
- Martin, C. D., Garcia, X., Breton, A., Thierry, G., & Costa, A. (2014). From literal meaning to veracity in two hundred milliseconds. *Frontiers in Human Neuroscience*, 8 (40), 1–12.
- Martin, C. D., Garcia, X., Breton, A., Thierry, G., & Costa, A. (2015). World knowledge integration during second language comprehension. *Language, Cognition and Neuroscience*, 31(2), 206–216.
- Master, P. (1987). *A cross-linguistic interlanguage analysis of the acquisition of the English article system*. (Doctoral Dissertation), University of California, Los Angeles.
- Master, P. (1997). The English article system: acquisition, function and pedagogy. *System*, 25(2), 215–232.
- Murphy, G. L. (1984). Establishing and accessing referents in discourse. *Memory & Cognition*, 12(5), 489–497.
- Murphy, S. L. (1997). *Knowledge and production of English articles by advanced second language learners*. (Doctoral dissertation), The University of Texas, Austin.

- Mitchell, D. C., & Green, D. W. (1978). The effects of context and content on immediate processing in reading. *The Quarterly Journal of Experimental Psychology*, 30(4), 609–636.
- Nakano, H., Saron, C., & Swaab, T. (2010). Speech and span: Working memory capacity impacts the use of animacy but not of world knowledge during spoken sentence comprehension. *Journal of Cognitive Neuroscience*, 22(12), 2886–2898.
- Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science*, 14, 11–28.
- Novick, J. M., Thompson-Schill, S. L., & Trueswell, J. C. (2008). Putting lexical constraints in context into the visual-world paradigm. *Cognition*, 107(3), 850-903.
- O'Grady, W. (2015). Processing Determinism. *Language Learning*, 65(1), 6–32.
- O'Grady, W., Kwak, H.-Y., Lee, O.-S., & Lee, M. (2011). An Emergentist perspective on Heritage Language Acquisition. *Studies in Second Language Acquisition*, 33(02), 223–245.
- Parrish, B. (1987). A new look at methodologies in the study of article acquisition of learners of ESL. *Language Learning*, 37(3), 361–383.
- Pérez-Leroux, A. T. (2011). What I don't understand about interfaces. *Linguistic Approaches to Bilingualism*, 1(1), 71–73.
- Phillips, C., & Ehrenhofer, L. (2015). The role of language processing in language acquisition. *Linguistic Approaches to Bilingualism*, 5(4), 409–453.
- Psychology Software Tools, Inc. [E-prime 2.0]. (2012). Retrieved from <http://www.pstnet.com>.
- Raudenbush, S., W., & Bryk, A., S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage publications.

- R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Robertson, D. (2000). Variability in the use of the English article system by Chinese learners of English. *Second Language Research*, 16(2), 135–172.
- Robinson, P. (2015). The cognition hypothesis, second language task demands, and the SSARC model of pedagogic task sequencing. In M. Bygate (Ed.), *Domains and directions in the development of TBLT* (pp. 87–122). Amsterdam/Philadelphia: John Benjamins.
- Russell, B. (1905). On denoting. *Mind*, 14(56), 479–493.
- Schultz, B. (2006). *Wh-scope marking in English interlanguage grammars: Transfer and processing effects on the second language acquisition of complex wh-questions*. [Doctoral dissertation], University of Hawai‘i at Mānoa, Honolulu, HI
- Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, 71, 109–147.
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, 10(1–4), 209–232.
- Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied linguistics*, 17(1), 38–62.
- Schumacher, P. B. (2009). Definiteness marking shows late effects during discourse processing: Evidence from ERPs. In S. L. Devi, A. Branco, & R. Mitkov (Eds.), *Anaphora processing and applications, Lecture notes in artificial intelligence* (Vol. 5847, pp. 91–106). Heidelberg: Springer.
- Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving

- incremental semantic interpretation through contextual representation. *Cognition*, 71, 109-147.
- Sheen, Y. (2007). The effects of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly*, 41(2), 255–283.
- Snape, N., Leung, Y.-k. I., & Ting, H.-C. (2006). Comparing Chinese, Japanese, and Spanish speakers in L2 English article acquisition: Evidence against the Fluctuation Hypothesis? In M. G. O'Brien, C. Shea, & J. Archibald (Eds.), *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference* (pp. 132–139). Somerville, MA: Cascadilla Proceedings Project.
- Snedeker, J., & Casserly, E. (2010). Is it all relative? Effects of prosodic boundaries on the comprehension and production of attachment ambiguities. *Language and Cognitive Processes*, 25(7-9), 1234-1264.
- Snedeker, J., & Trueswell, J. (2003). Using prosody to avoid ambiguity: effects of speaker awareness and referential context. *Journal of Memory and Language*, 48(1), 103–130.
- Sorace, A. (2011). Pinning down the concept of “interface” in bilingualism. *Linguistic Approaches to Bilingualism*, 1(1), 1–33.
- Sorace, A., & Filliaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second Language Research*, 22(3), 339–368.
- Sur, S., & Sinha, V. K. (2009). Event-related potential: An overview. *Industrial Psychiatry Journal*, 18(1), 70–73.
- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, New Series*, 268(5217), 1632-1634.

- Tarone, E. (1980). Communication strategies, foreigner talk, and repair in interlanguage. *Language Learning*, 30(2), 417–431.
- Tarone, E. (1983). On the variability of interlanguage systems. *Applied Linguistics*, 4(2), 142–163.
- Tarone, E., & Parrish, B. (1988). Task-related variation in interlanguage: The case of articles. *Language Learning*, 38(1), 21–44.
- Thomas, M. (1989). The acquisition of English articles by first- and second-language learners. *Applied Psycholinguistics*, 10(3), 335–355.
- Tremblay, A., & Demuth, K. (2007). Prosodic licensing of determiners in children's early French. In A. Belikova, L. Meroni, & M. Umeda (Eds.), *Proceedings of the 2nd conference on Generative Approaches to Language Acquisition North America (GALANA)* (pp. 426-436). Somerville, MA: Cascadia Proceedings Project.
- Trenkic, D. (2004). Definiteness in Serbian/Croatian/Bosnian and some implications for the general structure of the nominal phrase. *Lingua*, 114(11), 1401-1427.
- Trenkic, D. (2007). Variability in second language article production: beyond the representational deficit vs. processing constraints debate. *Second Language Research*, 23(3), 289–327.
- Trenkic, D. (2008). The representation of English articles in second language grammars: Determiners or adjectives? *Bilingualism: Language and Cognition*, 11(1), 1–18.
- Trenkic, D., Mirkovic, J., & Altmann, G. T. M. (2013). Real-time grammar processing by native and non-native speakers: Constructions unique to the second language. *Bilingualism: Language and Cognition*, 17(02), 237–257.

- Trenkic, D., & Pongpairaj, N. (2013). Referent salience affects second language article use. *Bilingualism: Language and Cognition*, 16(01), 152–166.
- Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: studying on-line sentence processing in young children. *Cognition*, 73, 89-134.
- Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, 33, 285-318.
- Tryzna, M. (2009). Questioning the validity of the Article Choice Parameter and the Fluctuation Hypothesis: Evidence from English article use by L1 Polish and L1 Mandarin speakers. In M. d. P. Garcia Mayo & R. Hawkins (Eds.), *Second Language Acquisition of Articles: Empirical findings and theoretical implications* (pp. 67–86). Amsterdam/Philadelphia: John Benjamins.
- Van Den Branden, K., Bygate, M., & Norris, J. (Eds.). (2009). *Task-based language teaching: introducing the reader*. Amsterdam, Netherlands: John Benjamins.
- VanPatten, B. (1990). Attending to form and content in the input. *Studies in second language acquisition*, 12(3), 287-301.
- VanPatten, B. (2002). Processing instruction: An update. *Language Learning*, 52(4), 755–803.
- Vanatten, B., Collopy, E., Price, J. E., Borst, S., & Qualin, A. (2013). Explicit Information, Grammatical Sensitivity, and the First-Noun Principle: A Cross-Linguistic Study in Processing Instruction. *The Modern Language Journal*, 97(2), 506-527.
- Wexler, K. (2011). Cues don't explain learning. In E. Gibson & N. J. Pearlmuter (Eds.), *The processing and acquisition of reference*. Cambridge, MA: The MIT press.